The International Congress for
global Science and Technology

ICGST International Journal on Graphics, Vision and Image Processing (GVIP)

Volume (15), Issue (II)
December 2015

www.icgst.com
www.icgst-amc.com
www.icgst-ees.com

GVIP Journal
ISSN Print 1687-398X
ISSN Online 1687-3998
ISSN CD-ROM 1687-4005

Table of Contents

ICGST International Journal on Graphics, Vision and Image Processing (GVIP)

A publication of the International Congress for global Science and Technology - (ICGST)

ICGST Editor in Chief: Dr. rer. nat. Ashraf Aboshosha

www.icgst.com, www.icgst-amc.com, www.icgst-ees.com

editor@icgst.com

www.icgst.com

Cellular Automata Clustering Based on Morphological Reconstruction (CACMR)

[1]G. Srinivasa Rao,[2]V.Vijaya Kumar, [3]Penmesta Suresh Varma
[1]*Pragathi Engineering College, kakinada, AP,India*
[2]*Anurag Group of Institutions, Hyderabad, Telangana,India*
3Adikavi Nannayya University,Rajahmundry,AP,India
[1]sr.goteti@gmail.com, [2]vakula_vijay@yahoo.com, [3]vermaps@yahoo.com

Abstract
A cellular automaton (CA) is one of the powerful and popular decentralized computing models with a wide variety of applications. The CA consists of an array of similar cells that interact with each another in a neighbourhood relationship and have definite state. The clustering and classifier models of CA are very popular in deriving significant knowledge from the large volumes of data set. The present paper presents a cellular automata clustering using morphological reconstruction (CACMR) on brain MRI images. The proposed morphological reconstruction operator segments accurately the skull and the brain as well as the background to accurately identify the tumour. For clustering the tumour the present paper uses an efficient cellular automata method based on Moore neighbourhood. The proposed CACMR is tested on brain MRI collected from various websites, and the various statistical measures that indicate the significance and accuracy of the proposed method.

Keywords:*Reconstruction, segmentation, MRI, Moore neighbourhood, statistical measures.*

1. Introduction
The decentralized computing model that is capable of providing an excellent platform for performing complex computation is the cellular automata (CA). The CA provides this with the help of only local information. The CA paradigm of local information, decentralized control and universal computation for modelling different applications are well established and exploited by the researchers, scientists and practitioners from various fields for the last decades. The reason behind the popularity of cellular automata can be traced to their simplicity, and to the enormous potential they hold in modelling complex systems, in spite of their simplicity. Cellular automata (CA) are arithmetical brands [4] and non-linear dynamical systems. In CA space and time are distinct and are termed as cellular, because they are made up of cells. Thesecells are treated as points in the lattice or square of the checker boards and are referred to as 'automata' [12]. The CAinclude a vast number of reasonably trouble-free individual units, or "cells". Each cell, in turn, is a straightforward predetermined automation which continually refreshes its own status, whereas the fresh cell state is dependent on the existing state of the cell and its immediate (local) neighbours [11,16] . CAs have been engaged extensively to assess intricate systems in environment [7], based on the above gleaming qualities. The provision of proficient and high-speed techniques without infringing the limitations of the data stream atmosphere has emerged as the most important challenge. This has ultimately and inevitably led to the origin of the procedure of data mining in cellular automata [14]. The data mining methods with CA are highly essential in several application domains such as online photo and video streaming services, economic analysis, concurrent manufacturing process control, search engines, spam filters, security, and medical services [12]. The impressive nature of the Cellular automata can be characterized as follows: [2]. 1. A cellular automaton is discrete time space. 2. Each and every cell comprises a number of restricted states. 3. The entire cells are located in the identical location. 4. Each and every cell is rationalized simultaneously. 5. The regulation in each and every locality is dependent on the value of the locality around its neighbours 6. The regulation for fresh value of each and every locality is also based on value of restricted number of preceding conditions. Two basic parameters characterize the CA. These two parameters are referred as number of states denoted by kand the second one specifies a radius of its neighbourhood [19] denoted as r. Each cell of CA vector endowed with 2 states and 1 radius. The relative regulations are known as transition rules [9]. The cellular automata methods in data mining applications [20]encapsulates input data d from a definite time interval of inspection, and transition rules are derived based on the application of data mining methods [3]. CAs are capable with several merits for modelling, together with their decentralized method, straightforward to the intricacy

rule, the association of form with task and model with procedure, the comparative easiness in visualizing the model outcomes, their elasticity, their vibrant technique, and also their kinship with geographical data systems and remotely sensed data. Of all, the most noteworthy quality is its effortlessness [1].

Today a very good quality of data is generated by various applications such as high-speed networking, finance logs, sensor networks, and web tracking [13]. The enormous amount of data thus collected from various sources is developed as an unrestricted data sequence arriving at the port of the system [17]. This leads to a problem of how to deal such a gigantic data stream [15]. There is a feast of methods designed for the mining of such items or models from data streams [8] using CA. Javier de Lope et al. [22] have jubilantly advocated a data clustering algorithm founded on the concept of deeming the individual data items as cells forming part of a one-dimensional cellular automaton. It integrates the insights into both social segregation brands rooted on cellular automata theory, where the data items themselves have the capability to travel freely in lattices, and also from ants clustering algorithms. Liliana Perez et al. [23] have legitimately launched the incorporation of the ABM with CA technique to successfully tackle modelling at both fine and large spatial scales. The distinct nature of CA facilitates incorporation with raster-based geospatial datasets in GIS, and is also advantageous during the course of modeling. A straightforward technique of edge recognition in accordance with the cellular automata by means of using digital images also proposed in the literature [24]. The recognition process is generally applicable to both monochromatic and colour images. Haijun Wang et al. [25] have heftily launched a cloud-based CA to characterize ambiguity proliferation.

Mathematical morphology [33] is an image transformation technique that locally modifies geometric features through set operations. It is a powerful tool with various applications, such as nonlinear image filtering, noise suppression, smoothing, shape recognition, shape reconstruction [28, 29], skeletonization [5, 6, 10, 21], texture segmentation [27], classification purpose [30] and medical image processing [31, 32]; and it is becoming very common in image processing. There are three prerequisites for the fuller realization of the potential of morphology: i. Complex processing combining various morphological operations (including other operations, such as discrete-time cellular neural networks, linear filtering , and area calculation) ii. Processing with large and complex structuring elements iii. High-speed (real-time) processing.

The present paper is organized as follows. The section 2 describes about proposed methodology, section 3 describes about experimental results and discussions and section 4 gives conclusions.

2. Methodology

The proposed method basically consists of two phases/modules. In the first phase skull striping is performed. For this the present paper used morphological reconstruction, which segments the skull in a better way than the existing methods, since it over comes the noise effect, smoothens the image and segregate the background pixels. In the second phase to cluster the region of interest cellular automata is used. By this brain tumours can be detected easily and accurately. The accuracy of the clustered image is measured by using statistical measures. The beauty of the proposed method is it applied morphological opening by reconstruction for segmenting of magnetic resonance image (MRI) of the brain. MRI is characterized for its high spatial resolution and soft tissue contrast. These two important characteristics make MRI as one of the most useful, significant and important imaging modalities in the diagnosis of brain related pathologies. The proposed morphological reconstruction by opening segments, as accurately as possible, the skull and the brain as well as background. In the proposed method elimination of skull from MRI is performed by reconstruction. The OTSU threshold on this enables the region of interest. Finally the proposed cellular automata clustering based on Moore neighborhood enables the region of interest in the MRI.

2.1 Proposed morphological reconstruction

The present paper used a new variant of morphological reconstruction method. Morphological reconstruction is very little-known method for extracting significant information about shapes in an image. The shapes could be just about anything: Skull of a MRI, letters in a scanned text document, fluorescently stained galaxies in a far-infrared telescope image etc... One can use morphological reconstruction in several applications and some of them are listed here. I) to extract marked objects ii) distinguish and remove objects that are touching the image border iii) filter out spurious high or low points iv) locate bright regions surrounded by shady or darker pixels v) distinguish or fill in object holes. There are two types of morphological reconstructions one is morphological Opening by Reconstruction and second one is morphological closing by reconstruction.

2.2 Opening by reconstruction

Opening is defined as erosion followed by dilation with a given structuring element on the image. The structuring element plays a vital role in this operation. In morphological opening, little or small objects that are smaller than the structuring element are typically removed by erosion, and the successive dilation tends to restore or ‘regrows’ the remaining objects by the same shape. The main problem with opening is the accurateness of this restoration depends on the similarity between the structuring element and the shape. The proposed morphological opening by

reconstruction method of this paper overcomes the above disadvantage of the simple opening. The proposed morphological opening by reconstruction preserves or restores the basic or original shapes of the objects after erosion.

Two images and a structuring element (instead of a single image and structuring element) are needed to perform a morphological reconstruction transformation. The first image, is the marker image, is the starting point for the transformation. The image that constrains the transformation is the mask image (second image). The structuring element defines connectivity between these. The proposed method used the original MRI as the mask image and the eroded MRI as the marker image. The proposed Morphological reconstruction filters enable the complete extraction of the marked objects by preserving the edges of the MRI. These filters preserve the contours. And also no new regional extrema are created by these reconstruction transformations on MRI. The fundamental mathematical morphology operations dilation and erosion are given in equation 1 and 2.

$$\text{Dilation}= D(A,B) = A \otimes B = \cup_{\beta\epsilon B}(A+B) \quad (1)$$

$$\text{Erosion} = E(A,B) = A\Theta\,(-B) = \cup_{\beta\epsilon B}(A-\beta) \quad (2)$$

Where A is original binary image and B is structure element.

The dilation operation fills the gaps or holes in the image and erosion widens the gaps.

The opening of image A by structuring element B, denoted, is defined as:

$$A \circ B = (A\Theta B) \oplus B \quad (3)$$

Thus, the opening A by B is the erosion of A by B, followed by a dilation of the result by B.

The closing of image A by structuring element B, denoted A • B, is defined as :

$$A \bullet B = (A \oplus B)\ \Theta\ B \quad (4)$$

The closing of a by B is simply the dilation of a by B, followed by the erosion of the result by B.

The structuring element B used in this paper is:

1	1	1
1	1	1
1	1	1

2.3 Clustering using cellular automata

The fundamental unit of cellular automata consists of a grid of cells and each cell has a finite number of states with finite number of dimensions. There are two type of neighborhood in CA; the first type is the von Neumann neighborhood in which each cell has four neighbors. The second one is the Moore neighborhood and it contains the eight cells i.e.it includes diagonal cells surrounded by a focal cell on a two-dimensional square grid. And it is named after Edward F. Moore, an initiator and pioneer of cellular automata hypothesis. Moore neighborhood is well utilized for the well-known Conway's Game of Life, which is based on the idea of 8-joined pixels in workstation representation. The idea might be stretched out to higher measurements. Therefore in the Moore neighborhood, the connected component (CC) in a tessellation (TT) is 8.

The proposed cellular clustering algorithm based on Moore neighborhood is given below.

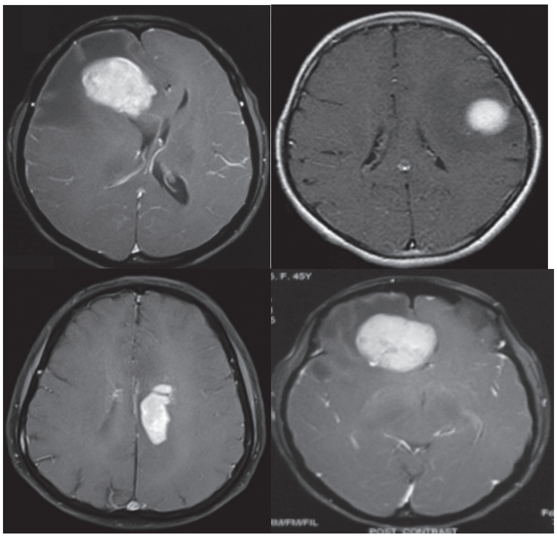
1. Obtain the resultant image from morphological reconstruction.
2. The boundary pixel is that pixel that differs with the current pixel.
3. Create a vector (V) and set as V = null
4. Examine the tessellation (TT) in all the direction to find whether any black pixel (Bp) (i.e. pixel value =0) exists in the tessellation (TT).
5. If any black pixel (Bp) exists in the tessellation, add it in to the vector (V).
6. Assign the found out black pixel (Bp) as the current boundary point (bp) and the pixel from which the black pixel (Bp) entered while the examining process is assigned as the neighborhood of the current boundary pixel (np).
7. Set the current pixel (cp) as the next clockwise pixel from the neighborhood of the boundary pixel (np).
8. Check whether the current pixel (cp) is equal to the black pixel (Bp) which is already detected.
9. If cp is a black pixel, then add it to the vector V and go to step 11.
10. If it is not then go to step 4.
11. Assign the neighborhood of the current boundary point (np) as the new boundary point (i.e, bp=np) .
12. Set the current pixel as the new neighborhood (i.e, np=cp).
13. Assign the current pixel (cp) as the next clockwise pixel from the neighborhood of the boundary pixel (np).
14. If cp is not a black pixel, then assign the current pixel (cp) as the next clockwise pixel from the neighborhood of the boundary pixel (np).
15. Update the neighborhood.
16. Continue this process until the current boundary pixel is equal to the boundary pixel for the second time.

Finally the clustered image is obtained. The termination condition given in step 16 restricts the set of shapes the calculation will walk totally.

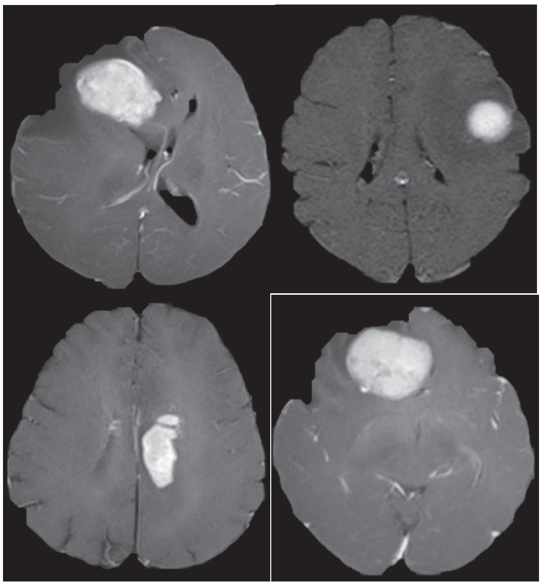
3. Experimental results and discussion

The proposed CACMR is experimented on a large scale (1000) of brain MRI, collected from various websites. The Figure 1(a) shows the four original MRI's, the Figure 1(b), 1(c) and 1(d) shows opening by reconstruction, image after threshold optimization

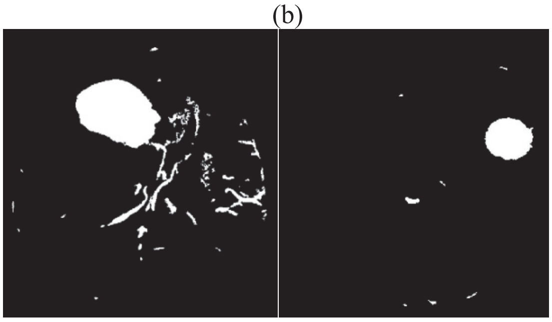
and-clustered image using cellular automata respectively for the corresponding original MRI's of Figure 1(a).

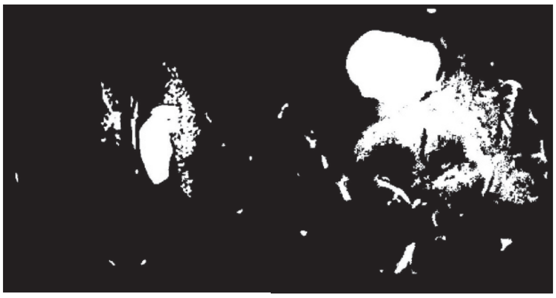


(a)

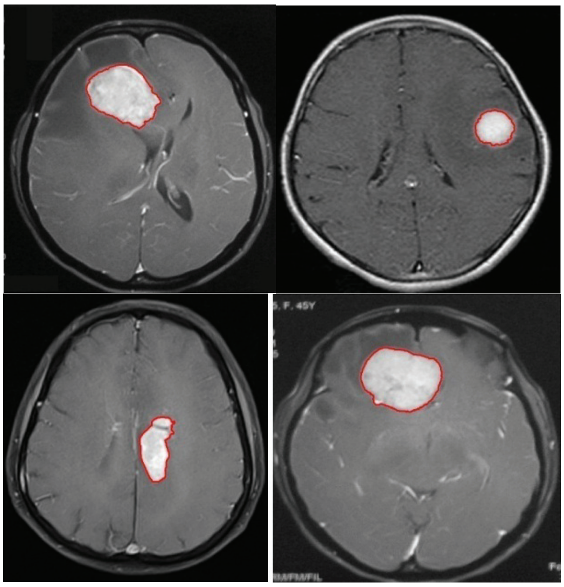


(b)





(c)



(d)

Figure 1: (a) Input image (b) Opening by reconstruction, (c) Image after threshold optimization and (d) Clustered image using cellular automata.

3.1 Performance Analysis

To evaluate the performance of the proposed CACMR accurately, the present paper evaluated statistical measures of the performance of a medical test such as sensitivity, specificity and accuracy. These parameters are widely quoted in statistics as a classification function. Sensitivity is also called as the true positive rate or the recall rate. The sensitivity parameter evaluates or measures the proportion of genuine or actual positives which are correctly identified. This clearly indicates the sensitivity is complementary to the false negative rate. Specificity (sometimes called the true negative rate) measures the proportion of negatives which are correctly identified as such (e.g., the percentage of healthy people who are correctly identified as not having the condition), and is complementary to the false positive rate. The average rate of the above parameters is plotted in the Figure 2. The proposed CACMR method outperforms the existing methods.

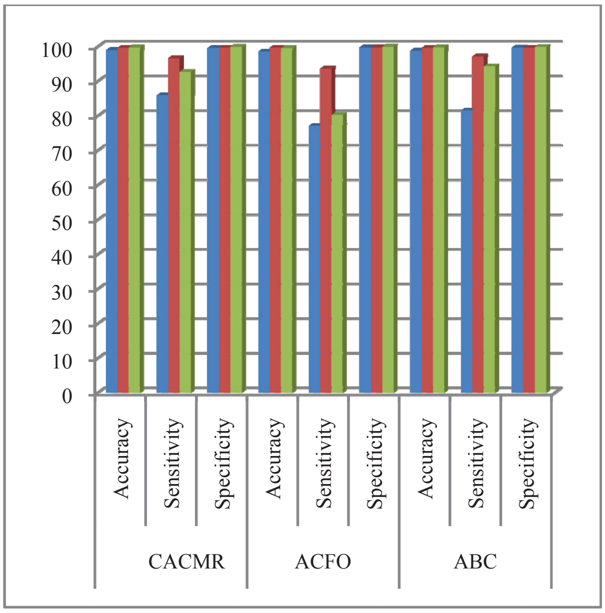


Figure 2: Performance measures comparison in terms of Accuracy, Sensitivity and specificity for the proposed technique and existing technique.

The present paper also evaluated false acceptance ratio (FAR) and false rejection ratio (FRR) on the above MRI by using the proposed and the existing methods and the results are displayed in the Figure 3. FAR and FRR are obtained by subtracting sensitivity and specificity values from one. Low value of FAR and FRR indicates better performance.

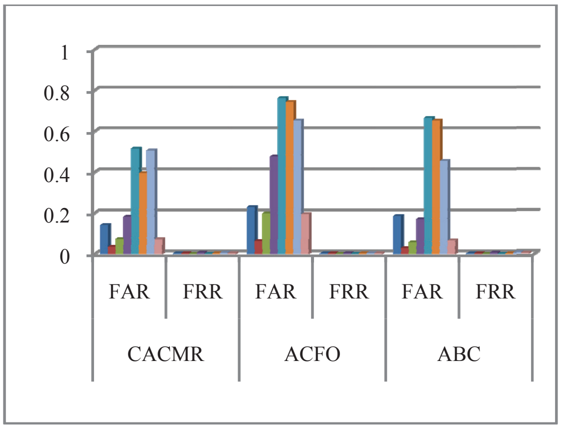


Figure 3: Performance measures comparison in terms of FAR and FRR for the proposed technique and existing technique.

The present paper also evaluated other performance measures like precision and recall to prove the efficacy. The Figure 4 indicates the average precision and recall rates. The average precision rate and recall of the proposed method is higher and lower respectively than the existing methods.

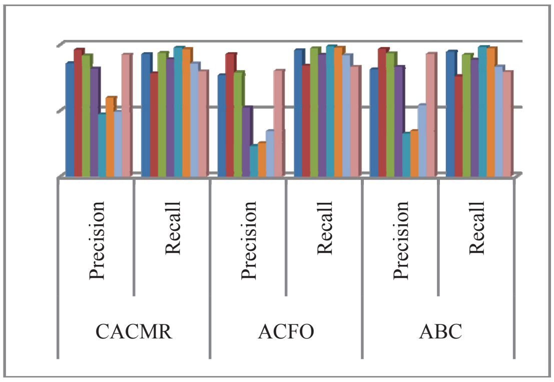


Figure 4: Performance of our proposed method and the existing technique in terms of Precision and Recall.

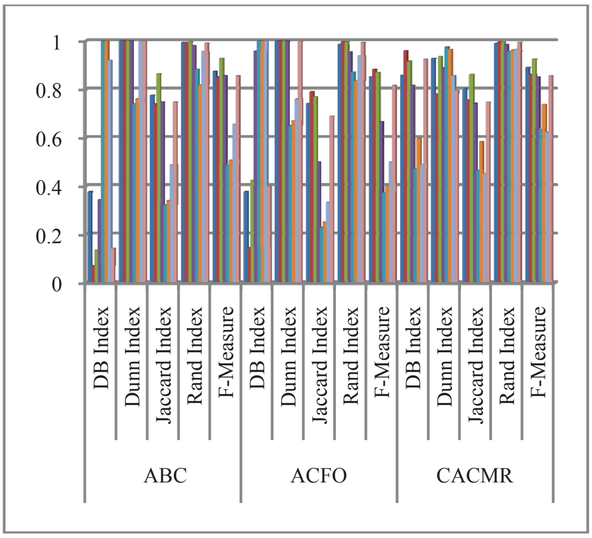


Figure 5: Performance comparison in terms of DB Index, Dunn Index, Jaccard Index, Rand Index and F-Measure for the proposed technique and existing technique.

3.2 Clustering Performance Evaluation by Benchmark Functions

To evaluate the performance of the proposed method the present paper evaluated the accurate internal evaluation schemes for clustering methods namely Davies–Bouldin index, Dunn Index, Jacard index, Rand index [26] and F-measure on the proposed CACMR and the other existing techniques. These internal evaluation schemes clearly specify how accurately the clustering has been done using the proposed quantities and feature of the derived model on the dataset. The clustering algorithm that produces a collection of clusters with the smallest Davies–Bouldin index is considered the best algorithm. High Dunn index value is more desirable for a clustering algorithm. The Jaccard index takes on a value between 0 and 1. An index of 1 means that the two datasets are identical and an index of 0 indicate that the datasets have no common elements. The Rand index is also viewed as a measure of the percentage of correct decisions made by the algorithm. F-measure combines both the precision and recall values into a single value as in figure 5.

When compared to the exiting technique, the proposed CACMR technique has low value of DB index, high values of Dunn index, Jaccard index, Rand index and F-measure which denotes that the proposed clustering technique is performs better than the existing technique.

4. Conclusion

The proposed cellular automata clustering efficiently used morphological reconstruction for segmenting the brain MRI. The main advantage of the proposed segmentation is to segment, as accurately as possible, the skull and the brain as well as the background. To eliminate the skull from the image, an opening by reconstruction of size 3 x 3 is applied. Some of the existing methods fail in segmenting if a thin connection exists between skull and brain. The proposed morphological reconstruction overcomes this. Also the performance of the proposed CACMR optimization technique is estimated using benchmark functions and the outcome proves that the proposed CACMR technique convergences faster than the existing technique.

References

[1] Mona Gamal, Ahmed Abou El-Fetouh and Shereef Barakat, "A Fuzzy Rough Rule Based System Enhanced By Fuzzy Cellular Automata", International Journal of Advanced Computer Science and Applications, Vol. 4, No. 5, pp. 1-11, 2013.

[2] SmitaNirkhi, Dharaskar and Thakre, "Data Mining: A Prospective Approach For Digital Forensics", International Journal of Data Mining & Knowledge Management Process, Vol. 2, No. 6, pp. 41-48, 2012.

[3] Pokkuluri KiranSree and Nedunuri Usha Devi, "Achieving Efficient File Compression with Linear Cellular Automata Pattern Classifier", International Journal of Hybrid Information Technology, Vol. 6, No. 2, pp. 15-26, 2013.

[4] Andrew Ninh, "Two Discrete Stochastic Cellular Automata Models of Cancer Stem Cell Proliferation", International Journal of Bioscience, Biochemistry and Bioinformatics, Vol. 3, No. 5, pp. 479-482, 2013.

[5] V.Vijaya Kumar, A. Srikrishna, Sadiq Ali, S. Trinath , A new skeletonization method based on connected component approach, International journal of computer science and network security (IJCSNS), Vol. 8, No.2 pp. 133-137, Feb-2008, ISSN: 1738-7906.

[6] V.Vijaya Kumar, A. Srikrishna, D.V.L.N Somayajulu, B. Raveendra Babu, An improved iterative morphological decomposition approach for image skeletonization, ICGST-Graphics, vision and image processing (ICGST-GVIP), Vol.8, No.1, pp.47-54, June 2008, ISSN 1687-398X..

[7] Mansour Esmaeilpour, Vahideh Naderifar and Zarina Shukur, "Cellular Learning Automata Approach for Data Classification", International Journal of Innovative Computing Information and Control, Vol. 8, No. 12, pp. 8063-8076, 2012.

[8] Amutha Mary and Dr. T. Jebarajan, Performance Metrics of Clustering Algorithm, Indian Journal of Applied Research, Vol. 4, No. 8, pp. 165-167, 2014

[9] Haijun Wang, Sanwei He, Xingjian Liu, Lan Dai, Peng Pan, Song Hong and Wenting Zhang, "Simulating urban expansion using a cloud-based cellular automata model", Landscape and Urban Planning, Vol. 110, pp. 99-112, 2013.

[10] V. Vijaya Kumar, U.S.N. Raju, P. Premchand, A. Suresh, Skeleton primitive extraction method on textures with different nonlinear wavelets, Journal of computer science, Vol.4, No.7, pp. 591-599, July-2008, ISSN: 1549-3636..

[11] Liliana Perez and Suzana Dragicevic, "Landscape level simulation of forest insect disturbance Coupling swarm intelligent agents with GIS-based cellular automata model", Ecological Modeling, Vol. 231, pp. 53–64, 2012.

[12] Javier de Lope and Dar Maravall, "Data clustering using a linear cellular automata based algorithm", Neuro computing, Vol. 114, No. 2, pp. 86–91, 2013.

[13] Fang Wang, Jean-Gabriel Hasbani, Xin Wang and Danielle Marceau, "Identifying dominant factors for the calibration of a land-use cellular automata model using Rough Set Theory", Computers Environment and Urban Systems, Vol. 35, No. 2, pp. 116-125, 2011.

[14] Tipawan Silwattananusarn and Kulthida Tuamsuk, "Data Mining and Its Applications for Knowledge Management", International Journal of Data Mining & Knowledge Management Process, Vol. 2, No. 5, pp. 13-24, 2012.

[15] Satinder Singh and JawaharLal, "Role of Cellular Automata in Budding Fields of Computer Science", International Journal of Advanced Computer and Mathematical Sciences, Vol. 3, No. 4, pp. 550-555, 2012.

[16] KiranSree and Ramesh Babu, "Identification of Promoter Region in Genomic DNA Using Cellular Automata Based Text Clustering", The International Arab Journal of Information Technology, Vol. 7, No. 1, pp. 75-78, 2010.

[17] BijayalaxmiKar, ChandrasekhraRao and Amiya Kumar Rath, "Generating PNS for Secret Key Cryptography Using Cellular Automaton", International Journal of

Advanced Computer Science and Applications, Vol. 2, No. 5, pp. 101-105, 2011.

[18] Karol Kozak, AagyaAgrawal, NikolausMachuy and Gabor Csucs, "Data Mining Techniques in High Content Screening: A Survey", Journal of Computer Science & Systems Biology, Vol. 2, No. pp. 219-239, 2009.

[19] SomanathTripathy and Sukumar Nandi, "Lightweight Cellular Automata-based Symmetric-key Encryption", International Journal of Network Security, Vol. 8, No. 2, pp. 243-252, 2009.

[20] HamouReda Mohamed, Lokbani Ahmed Chaouki, Ahmed Lehireche and Rahmani Mohamed, "Text Clustering Based on The N-Grams By Bio Inspired Method", Journal of Arts Science & Commerce, Vol. 1, No. 1, pp. 56-68, 2010.

[21] U.S.N. Raju, B. Eswar Reddy, V.Vijaya Kumar, B. Sujatha, "Extraction of skeleton primitives on wavelets, Journal of theoretical and applied information technology, Vol.4, No.11, pp.1065-1074, Nov- 2008, ISSN: 1992-8645.

[22] Milea, Teodorescu, Muller, Dragulinescu, Oltu, Tiplea, Stefan and Pompilian, "Cellular Automata Applications for Renewable Energy Monitoring", International Journal of Energy And Environment, Vol. 5, No. 3, pp. 418-425, 2011.

[23] RaminJavadzadeh, Zahra Afsahi and Mohammad Reza Meybodi, "Hybrid Model Based on Artificial Immune System and Cellular Automata", World Academy of Science, Engineering and Technology, Vol. 59, No. 1, pp. 658-661, 2011.

[24] Ahmad Zuhdi1, Aniati Murni Arymurthy and Heru Suhartanto, "Geographic Spatiotemporal Dynamic Model using Cellular Automata and Data Mining Techniques", International Journal of Computer Science Issues, Vol. 8, No. 2, pp. 53-61, 2011.

[25] Kiran Sree and Ramesh Babu, "Non Linear Cellular Automata in Predicting Heart Attack", International Journal of Hybrid Information Technology, Vol. 4, No. 1, pp. 33-40, 2011.

[26] Shanthi and Rajan, "Agent Based Cellular Automata: A Novel Approach for Modeling Spatiotemporal Growth Processes", International Journal of Application or Innovation in Engineering & Management, Vol. 1, No. 3, pp. 56-61, 2012.

[27] V. Vijaya Kumar, B. Eswar Reddy, A. Nagaraja Rao, U.S.N. Raju, “Texture segmentation methods based on combinatorial of morphological and statistical operations”, Journal of multimedia (JMM), Academy publishers, Vol.3, No.1, pp.36-40, May-2008, JMM, ISSN 1796-2048.

[28] V. Vijaya Kumar, N. GnaneswaraRao, A.L.NarsimhaRao, And V.VenkataKrishna,”IHBM: Integrated histogram bin matching for similarity measures of color image retrieval”, International journal of signal processing, image processing and pattern recognition, Vol. 2, No.3, pp. 109-120, Sep-2009, ISSN: 2005-4254.

[29] V. Vijaya Kumar, Srikrishna, “Error free iterative morphological decomposition algorithm for shape representation”, Journal of computer science, Vol.5, No.1, pp. 71-78, 2009, ISSN: 1549-3636.

[30] V Vijaya Kumar, U S N Raju, K Chandra Sekaran, V V Krishna, “Employing long linear patterns for texture classification relying on wavelets”, ICGST-Graphics, vision and image processing (ICGST-GVIP), Vol.8, No.5, pp. 13-21, Jan-2009, ISSN: 1687-398X.

[31] V. VijayaKumar, SakaKezia, I.SantiPrabha, “A new texture segmentation approach for medical images”, International journal of scientific & engineering research, Vol. 4, No.1, pp.1-5, January-2013, ISSN: 2229-5518.

[32]. M .Joseph Prakash, V. Vijayakumar, “A new texture based segmentation method to extract object from background”, Global journal of computer science and technology graphics & vision (GJCST), Vol.12, No.15, pp;1-6, 2012, ISSN: 0975-4350.

[33] J. Serra, Image Analysis and Mathematical Morphology. New York: Academic, 1982.

Biographies

Mr. Goteti Srinivasa Rao, obtained his Bachelor's degree in Computer Science and engineering from Ambedkar University, Maharashtra, India. He obtained his M.Tech degree in Computer Science and Engineering from JNT University, Kakinada, Andhra Pradesh,in the year 2004. He started his career at Polytechnic college level as a Lecturer. He Worked as Assistant professor in BVC college of Engineering, Amalapuram, AndhraPradesh.He worked as Associate Professor in Pragati Engineering college, Kakinada .His Specializations include Software Project management, Biometrics and Data mining..His Current research interest is in Data Mining with Cellular Automata.

Dr. VakulabharanamVijaya Kumar is working as Professor & Dean in Dept. of CSE & IT at Anurag Group of Institutions,(AGOI)(Autonomous), Hyderabad. He received integrated M.S.Engg, in CSE from USSR in 1989. He received his Ph.D. degree in Computer Science from Jawaharlal Nehru Technological University (JNTU), Hyderabad, India in 1998 and guided 18 research scholars for Ph.D. He has served JNT University for 13 years as Assistant Professor and Associate Professor. He is also acting as Director for Centre for Advanced Computational Research (CACR) at AGOI, Hyderabad where research scholars across the state are working. He has received best researcher and best teacher award from JNT University, Kakinada, India. His research interests include Image Processing, Pattern Recognition, Digital Water Marking, Cloud Computing and Image Retrieval Systems. He is the life member of CSI, ISCA, ISTE, IE (I), IETE, ACCS, CRSI, IRS and REDCROSS. He published more than 120 research publications till now in various National, International journals and conferences. He has also established and acted as a Head, SrinivasaRamanujan Research Forum (SRRF) at GIET, Rajahmundry, India from May 2006 to April 2013 for promoting research and social activities.

Dr.Penmesta.Sureshvarma is a profesor in Computer Science and working as Dean, College Development Council,AdikaviNannayya University Rajahmundry. He is also Dean Faculty of Engineering and Technology. He served as Principal, university colleges since 4th Dec.2009 to 13th june 2012. He also served as Head of the Department of CSE during the year 2008 to 2011. He secured his M.Tech degree in CSE with First Class from Andhra University, Visakhapatnam in the year 1999. He received his Ph.D Degree in Computer Science and Engineering from AcharyaNagarjuna University, Guntur in the year2008. He is honored with Best Teacher award by the Govt. of Andhra Pradesh in the year 2010 on the occasion of Teachers' Day. To his research credit, he has guided 4 Ph.D Scholars who were awarded their Ph.D Degrees. He also guided more than 100 M.Tech/MCA students in their thesis completion work. His areas of research interest include Communication Networks, Network Security, Data Mining and Image Processing. He is the Life Member of FIE, MIEEE, MASM, SMCSI, MISTE, SMORSI, MISCA, SMIA, CSIT, MIISA, MJAOE, MIAENG, AMIE and MAPSMS. He has published more than 100 research papers in various National and International journals.

Graphics, Vision and Image Processing, V. 15, No. 2, ISSN 1687-398X, Delaware, USA, December 2015

www.icgst.com

Heterogeneous Features using S-transform and Local Binary Patterns for Non-ideal Iris Recognition

P.V.L. Suvarchala, S. Srinivas Kumar and B. Chandra Mohan
JNTUK, Kakinada,Andhra Pradesh, India
suvarchala_pvl@yahoo.com

Abstract

Iris Recognition under non ideal imaging conditions like eyelash occlusions, rotation of eye and Charge Coupled Device (CCD) noise etc., is a challenging problem that sought the attention of researchers. In the proposed method heterogeneous features are extracted using 2D Discrete Orthonormal Stockwell Transform (DOST) and rotation invariant Local Binary Patterns (LBP). The DOST provides frequency information where as the LBP provides the spatial textural information. The feature set size is reduced based on the entropy of the features, which are used to train and test the Support Vector Machines (SVM) for checking the classification accuracy. The verification performance of the proposed scheme is validated using the benchmark databases and the % Correct Recognition Rate (CRR) is found to be more than (99%) under non ideal imaging conditions also.

Keywords:*Biometrics, Iris recognition, S Transform, LBP, Statistical Moments, SVMs.*

1 Introduction

Iris Recognition is considered to be the popular and sophisticated biometric technique. Iris is the annular portion between dark pupil and white sclera in the human eye which is an internal organ yet externally visible and has an extraordinary texture that is unique for each individual. The contraction and dilation of the iris muscles control the amount of light entering the eye. The textural features like freckles, ridges, corona and collarette region are same throughout the life time and do not change at all. Iris biometric is well accepted by the public as the imaging technique is non invasive and gained the attention of researchers as the iris patterns are stable and unique. Several international airports have established the iris scan systems to identify their passengers and facilitate quick processing. The artefacts in iris images such as poor brightness, low contrast, motion blur, specular reflections, eye rotation, direction of gaze, pupil contraction and dilation, eyelid and eyelash occlusions increase the false negative rate. The decision environment for iris recognition is influenced by many unfavourable conditions which are shown in Figure. 2 such as non ideal imaging conditions, rotation of eye and direction of gaze, CCD noise and particularly eyelash and eyelid occlusions which was established in [4] and [5].

Whilst majority of the researchers focussed on preprocessing of iris region like segmentation and noise reduction, new trends have been introduced in the area of processing and recognizing nonideal iris images. The comparison of existing iris recognition algorithms is given in Table.1 and a comprehensive survey can be found in [2]. The relationship between the within-class variability and between-class variability is the core issue in pattern recognition. The degrees of freedom of classes of pattern determine it. The separation or overlap among pattern classes affect the decidability index of pattern recognition algorithm. Preferably the within-class variability should be small and between-class variability large.

To achieve this, the feature selection plays a crucial role. Most iris recognition algorithms use either multi resolution transforms like Gabor or spatial domain features for iris feature extraction. A new feature extraction algorithm based on heterogeneous features, a combination of 2D DOST proposed in [6], which is a multi resolution transform and rotation invariant LBP proposed in [9] which is spatial domain technique, is presented in this paper that gives classification rates more than 99% when trained and tested using SVMs.

The iris region is first segmented and then unwrapped. After normalization of the iris region, the textural features are extracted using 2D DOST which decomposes the image in to different frequencies with bandwidths increasing dyadically. The rotational invariant LBP features along with DOST coefficients form the feature set pool as shown in the Figure. 1. Then this feature vectors in the pool are subjected to entropy based reduction which are then classified using SVMs. SVMs are being successfully used for

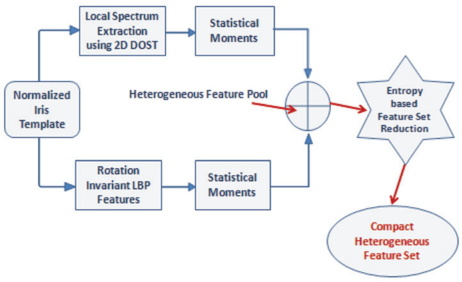


Figure 1: Iris Feature Extraction Algorithm

classification as they discriminate one class from the rest by a hyper plane with a great margin. The proposed method is validated using UBIRIS and CASIA, WVU and IITD databases.

The rest of the paper is organized as follows. In Section 2 the segmentation and normalization of iris region are explained. The heterogeneous textural feature collection and feature vector generation using DOST and LBP and feature set reduction based on entropy are presented in Section 3. The classification using SVMs is described in Section 4. In Section 5 the experimental results, comparisons and discussions are illustrated and in Section 6, the conclusions and scope of future work are presented.

1.1 Related Work

A detailed survey on iris recognition algorithms was presented in [2]. Zero crossing method with dissimilarity functions for matching was employed by Boles. The iris textural data were extracted using 2D Haar transform in [7]. Multi channel Gabor filtering for the extraction of iris features was used in [8]. Hilbert transform for iris feature extraction was used by Tisse. Levelsets were applied for segmentation and wavelet features were extracted in [10]. Circular Hough transform with dynamic range determination technique for improved segmentation of iris and pupil boundaries was proposed by Basma in [1]. A segmentation technique based on morphological processing which accurately detects the pupil boundaries was presented by Uma in [14]. Legendre moments were applied to capture the iris features corrupted by non ideal imaging conditions and CCD noise in [13]. Eyelash occlusions were eliminated using Monro Iris Transform algorithm in [17] in which the occluded pixels were filtered and the texture recovery was done.

In the proposed paper both 2D DOST coefficients and rotational invariant LBP features are extracted, their first order statistical moments are computed and then a compact feature vector set is built upon entropy based ranking. These feature vectors are then trained and tested using SVMs for checking the classification performance.

2 Iris Segmentation and Normalization

The circular iris and pupil regions are located by using Daugman's integro-differential operator. The dimensional inconsistencies amongst iris images are caused by contraction and dilation of pupil resulting from varying illumination levels during image acquisition. Also, the pupil region is not exactly concentric in the iris region. So for these reasons, the iris region is to be changed from circular to cartesian form of fixed dimensions, in order to facilitate comparison. These normalized iris images are enhanced using high boost filtering such that the edge details are emphasized as well as contrast is improved. The integro-differential operator for segmentation and rubber sheet model for normalization are given in [3].

3 Heterogeneous Feature Extraction using 2D DOST, LBP and Statistical Moments

The S-Transform named after the inventor Stockwell proposed in [12] suits well for the classification of noisy textures. It is proved that the S-Transform is more immune to noise than Wavelet transforms. The 2D DOST of an image is calculated in frequency domain using a dyadic sampling scheme. The 2D Fourier Transform of a discrete signal $f(x, y)$ in x and y directions is given by

$$F[m,n] = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f[x,y]\; e^{-2\pi i(\frac{mx}{M}+\frac{ny}{N})} \quad (1)$$

and the 2D inverse Fourier transform is given by

$$f[x,y] = \frac{1}{MN} \times \sum_{m=-M/2}^{M/2-1} \sum_{n=-N/2}^{N/2-1} F[m,n] e^{2\pi i(\frac{mx}{M}+\frac{ny}{N})} \quad (2)$$

The 2D DOST of an $N \times N$ image is computed by splitting the 2D FT of the image, $F[m, n]$, then multiplying with square root of the number of points in the block and then computing the inverse 2D-FT

$$S[x', y', v_x, v_y] = \frac{1}{\sqrt{2^{p_x+p_y-2}}} \sum_{m=-2^{p_x-2}}^{2^{p_x-2}-1} \sum_{n=-2^{p_y-2}}^{2^{p_y-2}-1} F[m+v_x, n+v_y] e^{2\pi i(\frac{mx'}{2^{p_x-1}}+\frac{ny'}{2^{p_y-1}})} \quad (3)$$

The Fourier spectrum is divided such that the wave numbers (v_{x_0}, v_{y_0}) are shifted to zero wave number

Table 1: Comparison of Proposed Method with Existing Iris Recognition Methods

Author	Segmentation Technique	Feature Set Nature	Matching Process	Quality Evaluation
Daugman [3],[4],[5]	Integro-differential operator	Binary feature vector using 2D Gabor filters	Hamming distance	Good recognition rate
Wildes et al. [16]	Image intensity gradient and Hough transform	Laplacian pyramid 1D signature	Normalized correlation	Matching process time consuming. Suitable for identification but not for recognition
Ma et al. [8]	Canny edge detection and Hough transform	1D real valued feature vector using wavelets	Weighted Euclidean distance	Local features used for iris recognition
Schuckers et al. [11]	Integro-differential operator and angular deformation method	ICA and Bi orthogonal wavelets	Hamming distance	Enhanced performance on non ideal data bases
Vatsa et al. [15]	Mumford shaw functional	1D gabor filter for textural features and Euler number for topological features	Iris indexing algorithm	Good identification rate on non ideal iris databases
Roy and Bhattacharya [10]	Levelset methods	Wavelet features	Adaptive Asymmetrical SVMs	High recognition rate with non ideal datasets
P.V.L. Suvarchala et al. [13]	Integro-differential operator	Exact Legendre moments	SVMs	High recognition rate in non ideal imaging conditions and CCD noise
Zhang et al. [17]		Monro Iris Transform	Weighted Hamming distance	Eyelash occlusions are removed using non linear filtering
Proposed Method	**Integro-differential operator**	**Heterogeneous features using 2D DOST and LBP**	**Support Vector Machines**	**Improved % CRR due to compact feature set**

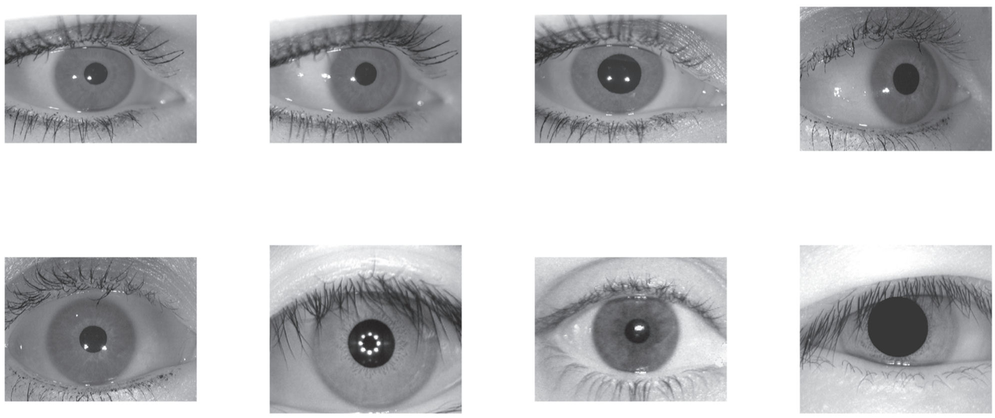


Figure 2: Iris images from CASIA, UBIRIS, WVU and IITD databases with specular reflections, pupil dilation, rotation of eye, eyelash occlusions, CCD noise, eyelid occlusion along with pupil dilation making very little active iris area

point and a $2^{p_x} - 1 \times 2^{p_y} - 1$ point inverse Fast Fourier Transform is applied which results in a rectangular image block of size $2^{p_x} - 1 \times 2^{p_y} - 1$. The total number of points in the 2D DOST image and the original image are the same.

The 2D DOST inverse can be found by applying forward 2D FFT to each image block to reverse the spectral partitioning such that the spectrum of the image is reconstructed.

$$F[m,n] = \sqrt{2^{p_x+p_y-2}} \sum_{m=-2^{p_x-2}}^{2^{p_x-2}-1} \sum_{n=-2^{p_y-2}}^{2^{p_y-2}-1} S[m - v_x, n - v_y]\ e^{-2\pi i (\frac{mx'}{2^{p_x}-1}' + \frac{ny'}{2^{p_y}-1}')} \tag{4}$$

Now the original image can be reconstructed by applying the inverse Fourier transform as follows:

$$f[x,y] = \frac{1}{N^2} \times \sum_{m=-N/2}^{N/2-1} \sum_{n=-N/2}^{N/2-1} F[m,n]\ e^{2\pi i (\frac{mx+ny}{N})} \tag{5}$$

The DOST and Discrete Wavelet Transform (DWT) both use a dyadic sampling scheme with the orders of $0, 1, 2, 4, ..., \log n - 1$. How ever they provide dissimilar information regarding the frequency content in the image. The DWT generates horizontal, vertical and diagonal detail coefficients for each order of sampling while the DOST provides the frequencies (v_x, v_y) with a bandwidth of $(2^{p_x} - 1, 2^{p_y} - 1)$.

DOST separates an image in to horizontal and vertical frequencies with different bandwidths. It can provide spatial frequency representation while maintaining the phase properties of the Fourier Transform. The additional feature of 2D DOST is, it can give pixel by pixel texture description of the image by generating the local spectrum that contains the horizontal and vertical frequency content from the Fourier Transform of the image given by [6].

The local spatial frequency description at a single pixel or patch of pixels can be calculated from $S[x', y', v_x, v_y]$ for all (v_x, v_y).

The local spectrum extraction algorithm proposed is as follows:

1. Resize the iris template $I(x,y)$ to 16×256 and normalize the intensity values to be between $0, 1$;
2. Divide the iris image into $16, 16 \times 16$ subimages $I_1, I_2, ..., I_{16}$;
3. Let horizontal and vertical voice frequencies be $v_x = 16$ and $v_y = 256$;
4. Create and initialize local spectrum matrix of size $log_2(v_x) \times log_2(v_y)$ to contain zeros;
5. Create and initialize feature vector FV to contain zeros;
6. **forall** sub images of $I_1(x,y)$ to $I_{16}(x,y)$ **do**
7. **forall** v_x in 1 to $log_2(v_x)$, and v_y in 1 to $log_2(v_y)$ **do**
8. $S[x, y, v_x, v_y] \leftarrow$ Compute 2D DOST;
9. **end**
10. Compute I order statistical moments for the local spectrum $S[x, y, v_x, v_y]$;
11. FV $\leftarrow$ statmoments;
12. **end**

3.1 Rotation Invariant LBP Features

LBP is an operator based on description of the signs of differences between the neighboring pixels in the image. It is immune to monotonic changes in the gray levels of the image. Gray scale and rotation inavariant LBP were introduced in [9] which characterize the local texture of the image and also the contrast. The gray scale invariant LBP in 3×3 neighborhood of a pixel is achieved using the expression as follows:

$$LBP_8 = \sum_{i=1}^{g} s(g_i - g_0)\ 2^{i-1} \tag{6}$$

where,

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \tag{7}$$

and g_0 is the center pixel gray level value and g_i, $i = 1, 2, ...8$ are the gray level values of eight surrounding pixels of circularly symmetric neighborhood as shown in Figure. 3. Rotational invariance of LBP was developed based on the assumption that, out of possible $2^8 = 256$ patterns only 36 patterns are unique and in these also only 9 have uniformity measure (U) at most 2, which corresponds to spatial transitions in the

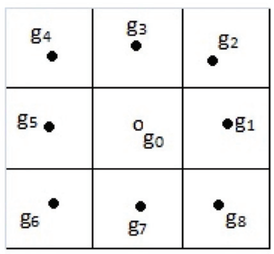


Figure 3: Circularly symmetric 3×3neighborhood

patterns (0/1 or 1/0). Such patterns are designated as uniform. Hence the rotational invariant LBP is expressed as

$$LBP_8^{riu2} = \begin{cases} \sum_{i=1}^{8} s(g_i - g_0) & \text{if } U(LBP_8) \leq 2 \\ 9 & \text{otherwise} \end{cases} \tag{8}$$

The LBP extraction algorithm proposed is as follows:

1. Resize the iris template $I(x, y)$ to 16×256;
2. Divide the iris image into $16, 16 \times 16$ subimages $I_1, I_2, ..., I_{16}$;
3. Create and initialize lbpmat of size of subimage to contain zeros;
4. Create and initialize feature vector FV to contain zeros;
5. **forall** sub images of $I_i(x, y)$ with i in 1 to 16 **do**
6. *lbpmat* $\leftarrow$ Compute rotational invariant LBP$\{I_i(x, y)\}$;
7. Compute I order statistical moments for the lbpmat;
8. FV $\leftarrow$ statmoments;
9. **end**

3.2 Statistical Moments

If Ω_i be the i^{th} local DOST spectrum or local binary pattern consisting of $M \times N$ coefficients, the first order statistical moments, mean, standard deviation, skew and kurtosis are calculated for each block as follows: The first moment about origin is known as **mean** (μ_i) which gives the average value of the local spectrum or pattern coefficients in the i^{th} region:

$$mean(\mu_i) = \frac{1}{MN} \sum_{x1=1}^{M} \sum_{x2=1}^{N} I_{\Omega_i}(x1, x2) \tag{9}$$

The square root of second central moment is known as **standard deviation** (σ_i) is the measure of variability of the coefficients' values about the mean μ_i:

$$std(\sigma_i) = \sqrt{\frac{1}{MN} \sum_{x1=1}^{M} \sum_{x2=1}^{N} (I_{\Omega_i}(x1, x2) - \mu_i)} \tag{10}$$

The third central moment is known as **skew**, which describes how symmetric the coefficients are about its mean.

$$skew(s_i) = \frac{1}{MN\sigma_i^2} \sum_{x1=1}^{M} \sum_{x2=1}^{N} (I_{\Omega_i}(x1, x2) - \mu_i)^3 \tag{11}$$

The fourth central moment is known as **kurtosis** which is the measure of flatness of the coefficient values in the region

$$kurtosis(k_i) = \frac{1}{MN\sigma_i^4} \sum_{x1=1}^{M} \sum_{x2=1}^{N} (I_{\Omega_i}(x1, x2) - \mu_i)^4 - 3 \tag{12}$$

After finding out the local spectrum DOST coefficients, as well as rotation invariant LBP of each block in the iris template the first order moments are computed and appended to form the feature vector. At the end of each feature vector the corresponding class label is attached, facilitating for the supervised learning and testing.

3.3 Entropy Based Feature Set Reduction

Before classification of the heterogeneous features of the iris images, ranking is given to the features according to their entropy. The entropy of a pattern which is in an order is lower than the entropy of the pattern with disorder. An irrelevant feature or signature has more randomness, which means uncertainty and hence higher entropy than a relevant signature. The feature with least entropy is given first rank and so on. The entropy is computed as follows.

$$\begin{aligned} Entropy = -\sum_{i=1}^{N} \sum_{j=1}^{N} d_{i,j} \times \log(d_{i,j}) + \\ (1 - d_{i,j}) \times \log(1 - d_{i,j}) \end{aligned} \tag{13}$$

where $d_{i,j}$ is the distance between features of N different samples in the same class. The features that have least entropy are more significant for classification. Roy in [10] proposed genetic algorithm based feature set reduction where as entropy based feature set reduction is computationally less complex than that and efficient.

4 Iris Classification with SVMs

In iris recognition many researchers applied Hamming distance for pattern matching [2]. Conventional SVMs and Non symmetrical SVMs were employed by a few researchers to distinguish the false positive and false negative cases and Asymmetrical Adaptive SVMs were used in [10] to cut the matching times of the test samples.

SVMs have been an excellent tool for data classification. The essential idea is to map the data points in to a high dimensional space and separating them by a hyper plane with highest margin. In the proposed method the heterogeneous feature vectors are trained and tested using traditional SVMs with Linear kernel at different train versus test ratios.

Table 2: Result on CASIA database - Comparison of EER($\times 10^{-3}$)

Method/ Parameter	Daugman	Tan	Monro	Roy	**Proposed Method**
EER	6.91	3.13	0.26	0.22	**0.18**

Table 3: EER of Proposed Method on Different Iris Databases at Train : Test Ratio of 70% : 30%

Iris Database	DOST Features-EER($\times 10^{-3}$)	LBP Features-EER($\times 10^{-3}$)	Heterogeneous Features-EER($\times 10^{-3}$)
CASIA-1	0.22	0.28	0.18
CASIA-Interval	0.96	1.2	0.82
UBIRIS-1	0.186	0.198	0.142
IITD	0.96	0.98	0.86
WVU	1.2	1.44	0.98

4.1 Classification Evaluation

The classification of iris feature vectors is carried out using osusvm package available at http://kaz.dl.sourceforge.net/project/svm/svm/3.00/osusvm-3.0.zip. The proposed method is tested at 70% : 30% Train:Test ratio with DOST features, rotation invariant LBP features independently as well as heterogeneous features using CASIA-1, CASIA-Interval, UBIRIS-1, WVU and IITD databases. The multi class SVMs operated in One-against-All approach are tuned with linear support vector classifier kernel. The performance of the proposed method is appraised in terms of Equal Error Rate (EER) and % CRR and compared with existing methods.

EER is the point at which the False Positive Rate (FPR) is equal to the False Negative Rate (FNR).

% CRR is defined as:

$$\%CRR = \frac{Correctly\ Recognized\ Users\ Number}{Total\ Number\ of\ Users\ Enrolled} \times 100 \tag{14}$$

5 Results of Experiments

The experiments are conducted on a PC with Pentium i3, 2 GHz processor and 2 GB RAM in MATLAB 7.10 environment. The evaluation of the proposed method is done on CASIA version-1, CASIA Iris Interval, UBIRIS version-1, WVU and IITD databases. The proposed method is checked on original as well as synthetic datasets of CASIA-1 and UBIRIS-1 generated under nonideal imaging conditions and CCD noise as described in [13].

CASIA version-1 and Iris Interval databases available at http://nlpr-web.ia.ac.cn/english/irds/ iris-database.html, comprise of 756 iris images taken in two sessions from 108 classes and 249 classes with left and right eyes respectively. Each iris image is of 8-bit gray scale and a resolution of 320×280. The UBIRIS version-1 database available at http://iris.di.ubi.pt, adds 1205 iris images from 241 subjects with 5 samples from each subject with a resolution of 150×200. The IITD iris database contributes 2220 images of 5 samples from left and right eyes of each class with a resolution of 320×240. The WVU offers a database of size 1852 images from 380 different people.

CASIA and IITD databases are of varying image quality, suffering from eyelash, eyelid occlusions where as UBIRIS database suffers from blur, specular reflections and partial iris also. WVU dataset images suffer from eye rotation and gaze rotation. Hence the four databases are preferred for the experiment as they contain the conditions suitable to evaluate the performance of the proposed work.

5.1 Discussion on Results

The iris region is segmented and normalized to obtain a 20×240 size iris template. The pre processing step is performed before extraction of DOST and LBP features, to enhance the contrast and highlight the edges in the image by applying high boost filtering. The spatio frequency domain and spatial domain features from all sub blocks are extracted and the coefficients are organized into a row forming a feature vector of size 128. After performing the step of entropy based ranking on the features, only 56 features with low entropy are considered for classification. Then feature vectors are classified using SVMs tuned with linear kernel.

It is found that the proposed work is compared with existing methods in terms of EER in Table 2. and the EER is very less compared to all. The heterogeneous feature set followed by entropy based ranking is of small size (56). The comparison of EER and % CRR is shown in Tables 3 and 4. Both EER and % CRR are improved in all datasets when both

Table 4: % CRR of Proposed Method on Different Iris Databases at Train : Test Ratio of 70% : 30%

Iris Database	DOST Features (%CRR)	LBP Features (%CRR)	Heterogeneous Features (%CRR)
CASIA-1	98.85	98.45	100
CASIA-Interval	97.4	97	99
UBIRIS-1	99.42	98.78	100
IITD	98.8	98.2	99.81
WVU	98.2	97.6	99.2

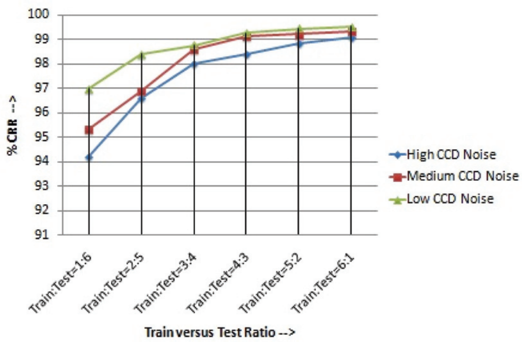


Figure 4: % CRR result on CASIA synthetic dataset - Non ideal imaging conditions and CCD noise simulated as mentioned in [13]

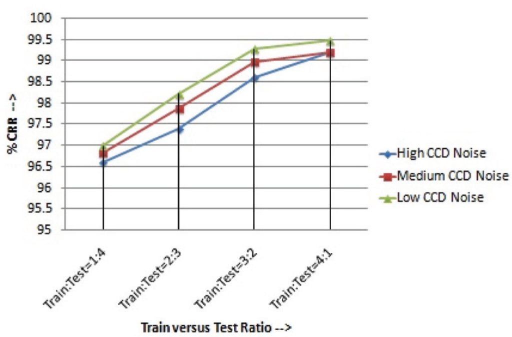


Figure 5: % CRR result on UBIRIS synthetic dataset - Non ideal imaging conditions and CCD noise simulated as mentioned in [13]

DOST and rotation invariant LBP features are combined rather than individually.

The robustness of proposed method is also evaluated on synthetic datasets of CASIA-1 and UBIRIS-1 in which, all non ideal imaging conditions like camera blur, low brightness, poor contrast and CCD noise are simulated as proposed in [13]. The evaluation is carried out at different train versus test ratios under low, medium and high variance of CCD noise and non ideal imaging conditions. The % CRR result on the synthetic datasets is shown in Figures. 4 and 5. It is observed that the proposed heterogeneous feature extraction method performed well under highly noisy conditions also.

6 Conclusion

A new feature extraction method using 2D DOST and rotation invariant LBP and statistical moments which is followed by entropy based ranking is proposed in this paper and it is proved that the performance of the iris recognition method is improved over the existing methods. It is proved from the results that, when the combined spatio frequency domain (DOST) features and spatial (rotation invariant LBP) domain features offer high recognition rates than independently. The heterogeneous features proved to be good on synthetic datasets also which suffer from non ideal imaging conditions and CCD noise. The feature vector size is very small (56) due to ranking and there is a scope for further reduction in size of the feature vector with feature set reduction techniques.

Acknowledgements

UBIRIS version-1, data set is released by department of computer science, University of Beira Interior, Portugal. CASIA version-1, data set is owned by Institute of Automation, Chinese Academy of Sciences, China, WVU dataset by West Virginia University, USA and IITD dataset by IIT Delhi, India.

References

[1] Basma M. Almezgagi, M.A. Wahby Shalaby, and Hesham N. Elmahdy. Improved Iris Verification System. *GVIP*, 14(1):29–38, 2014.

[2] K.W. Bowyer, Hollingsworth K., and Flynn P.J. Image Understanding for Iris Biometrics: A Survey. *Computer Vision and Image Understanding*, 10, 2008.

[3] J.G. Daugman. High Confidence Visual Recognition of Persons by a Test of Statistical Independence. *IEEE Trans. Pattern analysis and Machine Intelligence*, 15(11):1148–1161, 1993.

[4] J.G. Daugman. How Iris Recognition Works. *IEEE Transactions on Circuits, Systems and Video Technology*, 14, 2004.

[5] J.G. Daugman. New Methods in Iris Recognition. *IEEE Transactions Sys. Man and Cyber*, 37, 2007.

[6] Sylvia Drabycz, Robert G. Stockwell, and J.Ross Mitchell. Image texture characterization using discrete orthonormal S-transform. *Journal of Digital Imaging*, 22, 2009.

[7] S. Lim, K. Lee, O. Byeon, and T. Kim. Efficient Iris Recognition through Improvement of Feature Vector and Classifier. *IEEE Transactions On Image Processing*, 23, 2001.

[8] L. Ma, Y. Wang, and T. Tan. Iris Recognition Based on Multichannel Gabor Filtering. In *Fifth Asian Conf. Computer Vision.*

[9] Timo Ojala, Matti Pietikainen, and Topi Maenpaa. Gray scale and rotation invariant texture classification with local binary patterns. *IEEE Cong. Evol. Comp*, 22, 2006.

[10] K. Roy, Bhattacharya P, and Ching Y. Suen. Towards Nonideal Iris Recognition Based on Levelset Method, Genetic Algorithms and Adaptive Asymmetrical SVMs. *Engineering Applications of Artificial Intelligence*, 24, 2010.

[11] A. C. Stephanie Schuckers, Natalia A. Schmid, Aditya Abhayankar, Vivekanand Dorairaj, Christopher K. Boyce, and Lawrence A. Hornak. On Techniques for Angle Compensation in Nonideal Iris Recognition. *IEEE Trabsactions on Systems, Man, and Cybernetics-Part B: Cybernetics*, 37(5), 2007.

[12] R.G. Stockwell, L. Mansinha, and R. P. Lowe. Localization of the complex spectrum: The S transform. *IEEE Transactions on Signal Processing*, 44, 1996.

[13] P.V.L. Suvarchala, S. Srinivas Kumar, and B.Chandra Mohan. Iris Recognition under Non-ideal Imaging Conditions and CCD Noise. In *PReMI-2013, LNCS 8251, Springer-Verlag Berlin Heidelberg*, ISI, Kolkata, India, 2013.

[14] S. UmaMaheswari, P. Anbalagan, and T. Priya. Improved Iris Verification System. *GVIP*, 8(2):29–35, 2008.

[15] Mayank Vatsa, Richa Singh, and Afzel Noore. Improving Iris Recognition Performance Using Segmentation, Quality Enhancement, Match Score Fusion, and Indexing. *IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics*, 2008.

[16] R. Wildes, J. Asmuth, G. Green, S. Hsu, R. Kolcznski, J. Matey, and S. McBride. A Machine-Vision System for Iris Recognition. *Machine Vision and Applications*, 9, 1996.

[17] D. Zhang, D. M. Monro, and S. Rakshit. Eyelash removal method for human iris recognition. In *Image Processing, IEEE International Conference*, 2006.

Biographies

P.V.L.Suvarchala Author One received her AMIE from IEI and her M.Tech from JNTUH. Currently she is working for Amrita Sai Institute of Science and Technology and she is a PhD scholar at JNTUK, Kakinada, India, under the supervision of Dr. S. Srinivas Kumar. Her research interests include Image Processing and Pattern Recognition.

S.Srinivas Kumar Author Two is currently Professor of Electronics and Communication Engineering and Director, Research and Development Cell, JNTUK, Kakinada, India. He has been doing his research work in Digital Image and Video Processing for 20 years. His main research interests are Signal and Image Processing, Control Systems, Pattern Recognition, Machine Vision and Artificial Intelligence. email: samay_ssk2@yahoo.com

B. Chandra Mohan Author Three is currently Professor of Electronics and Communication Engineering and Dean, Academics at BEC, Bapatla, affiliated to ANU, India. He has been doing his research work in Digital Image and Video Processing for 10 years. His main research interests are Signal, Image and Video processing, Machine Vision and Artificial Intelligence. email: chandrabhuma@gmail.com

Graphics, Vision and Image Processing, V. 15, No. 2, ISSN 1687-398X, Delaware, USA, December 2015

www.icgst.com

PERFORMANCE ANALYSIS OF SINGLE DICTIONARY LEARNING FOR SINGLE IMAGE SUPER-RESOLUTION

Kiran Jadhav[1], Ramesh Kulkarni[2], Gaurav Tawde[3]
Department of Electronics and Telecommunication, V.E.S Institute of Technology, Mumbai, India
[1]kjkiranjadhav7@gmail.com, [2]ramesh.kulkarni@ves.ac.in, [3]gaurav.tawde@ves.ac.in

Abstract
In this paper, we propose a single dictionary learning algorithm to fully make use of only high-resolution images. Unlike the other methods, dictionary is trained from the set of high-resolution image patches of size 5x5, 7x7 and 9x9 instead of patch pairs of high-/low-resolution images. The advantage of this modification is no need to train the dictionary again when up-scaling is changed. There is a run-time improvement is achieved with best quality of reconstructed image due to single dictionary. The simulation results justify that the proposed method accomplishes the state-of-the-art results compared to other super-resolution methods in terms of both reconstruction ability and with shorter run-time. The demonstration results for single image super-resolution are more promising.

Keywords: *Bicubic Interpolation, Non Local Means, Super-resolution, Sparse Representation, Sparse Coding, Single Dictionary.*

1. Introduction
High-resolution images are useful in many applications, like in medical image diagnosis, remote sensing, video surveillance and satellite imaging [1]. But due to the technological and economic constraints, there is a restrain for availability of high-resolution images. Compared to other methods, super-resolution method is one of the best method for reconstructing high-resolution images from corresponding low-resolution images. In recent year, super-resolution reconstruction is an active area in image processing field which can help to overcome the resolution limitations of low-cost imaging sensors.

"Single image super-resolution" deals with the generating high-resolution (HR) image from single low-resolution (LR) image of the same scene. The most challenging task is to describe the relation between low- and high-resolution images [2], because over-complete sparse model are well developed to natural images and highly robust to the noise [3], based on learned dictionary, sparse representation has achieved best performance on image denoising and restoration [4].

Super-resolution task is the inverse problem which can be extremely ill-posed in nature. Because of many high-resolution images generated low-resolution images from same scene. In order to stabilize this ill-posed problem, regularization is an important procedure [5-7]. Super-resolution can be broadly classified into three major categories: interpolation, reconstruction and learning based super-resolution algorithm. Interpolation based super-resolution algorithm [8-10] are simple that generates smooth images and tend to blur high frequency details. In learning-based super-resolution algorithms [8], [11-13], detailed textures are explicated by seeking through a training set of low- and high-resolution images. They need a proper choice of the training images, otherwise inaccurate information may be found. Alternatively, availability of low-resolution and up-scaling factor is restricted in reconstruction-based super-resolution algorithms [14-19], apply constraints to the high-resolution images based on priori should generate original low-resolution images. Yang *et al*. in [12] developed sparse representation model to successfully reconstruct the high-resolution image patch. From a single input image, this method can reproduce both textures and precise edges. However, sparse representation is effectively used in many fields, like image compression, image denoising [20], face recognition [21] etc. The dictionary takes weighty role in sparse representation. Many dictionary learning methods has been developed to learn a dictionary from example patches [22-30].

The major limitation of these joint dictionary methods is that when the image is up-scaled, the dictionary is retrained. In this paper, we present a simple and efficient single image super-resolution algorithm to solve above problem. The advantages of this algorithm:

(1) The dictionary is trained from the high-resolution image patches only and hence, retraining of a new dictionary is not important when scaling factor changed. In addition, the training time is also reduces.

(2) Definitely run-time also reduces. We combine the both sparse representation patch-based

method and reconstruction-based method into a unified energy-function framework.

In our setting, we will work with only single dictionary training for high-resolution image patches.

This paper is organized as follows: Section (2) describes the mathematical background of related work, Section (3) represents the single dictionary learning and proposed algorithm, Section (4) shows experimental results and analysis of quality parameter and Section (5) concludes the paper.

2. Mathematical Background [31]

2.1 Non-Local Means

In image processing, the best algorithm for image denoising is Non-local means (NLM) filtering [32] which takes a mean of all pixels in the image, weighted by how similar these pixels are to the target pixel. This method has become more famous in many applications such as de-interlacing and view interpolation. This can be considered as a best kind of learning-based hallucination methods because training samples are originated from image itself in this method. Basically, NLM is a weighted filter which is formulated by

$$X_{ij} = \frac{\sum_{(p,q)\in P_{ij}} w_{ij}^{pq} Y_{pq}}{\sum_{(p,q)\in P_{ij}} w_{ij}^{pq}}, \tag{1}$$

Where Y_{pq} denotes the intensity of the low-resolution image Y at position (p, q) and x_{ij} denotes the intensity of high-resolution image X at position (i, j). P_{ij} is the index set including coordination of similar pixel with pixel (i, j). The weight W_{ij}^{pq} denotes the similarity between the patches $R_{ij}Y$ (the patch centered at (i, j) in image Y) and $R_{pq}Y$, which can be calculated by

$$w_{ij}^{pq} = \exp\left(-\frac{\left\|R_{ij}Y - R_{pq}Y\right\|_G^2}{h^2}\right), \tag{2}$$

where h play a role as a global smoothing parameter, parameter G is a kernel matrix that assigns the pixels close to the center of the patch to contribute more within target image patch and R_{ij} is an operator originating a patch centered at (i, j).

2.2 Sparse Representation

In signal processing, sparse representation refers as the sparse linear combination of small number of known or unknown "basis vectors". Basically sparse representation comprises of two things: (1) dictionary training and (2) sparse coding. Dictionary training is a new approach to learn a basis matrix D^{pxk} (every column a p-dimensional basis vector). The main goal of the training methods leads to simple formulation of l_0 and l_1 sparity measures.

Sparse coding is referred as the problem of finding sparse representation with a small number of significant coefficients. For sparse coding, consider signal x of dimension p can be sparsely represented with respect to a given dictionary D. This problem can be formulated by solving sparsest coefficient vector α as:

$$\min\|\alpha\|_0 \, s.t..x = D\alpha, \tag{3}$$

The single dictionary training process will be described in Section 3. In this setting, we use generic images for training purpose.

3. Single Dictionary Learning and Proposed Algorithm

This section will discuss the training of single dictionary. The dictionary is prepared only by set of high-resolution images.

In this technique, high-resolution image 'X' is reconstructed with low-resolution image 'Y'. In this proposed algorithm, three constraints are assumed to reconstruct high-resolution image X:

(1) Fidelity constrain: The reconstructed high-resolution image X should be consistent with the input low-resolution image Y according to image degradation model,

$$Y = SBX + \eta, \tag{4}$$

Where, Y is an observed low-resolution image which is down-sampled and blurred version of X, S denotes down-sampling operator, B denotes unknown blurring operator and η denotes noise item.

(2) Sparsity constrain: With respect to assumption of sparsity described in last section, reconstructed high-resolution image X can be spasrely represented with respect to unknown dictionary D^{pxk}.

$$\begin{gathered} D,\{\alpha_{ij}\} = \arg\min_{D,\{\alpha_{ij}\}}\left\|\alpha_{ij}\right\|_0 \\ s.t.\left\|R_{ij}X - D\alpha_{ij}\right\|_2^2 \le \varepsilon_{ij}, \end{gathered} \tag{5}$$

Where α_{ij} is a representation coefficient for patch $R_{ij}X$ and ε_{ij} denotes error item.

(3) Non-Local Mean constrain: This term shows that the one pixel in X could be estimated by a weighted average of the same pixel found in X. Non-local means regularization term for individual pixel x_{ij} can be formulated as

$$X_{ij} = \arg\min_{X_{ij}}\left\|X_{ij} - W_{ij}^T.G_{ij}\right\|_2^2, \tag{6}$$

Where x_{ij} denotes vector consists of similar pixels found in X and w_{ij} is the corresponding weight vector.

A unified energy-function framework is produced by combining all these three constraints together which can be written as:

$$D, X, \{\alpha_{ij}\} = \underset{D, X, \{\alpha_{ij}\}}{\arg\min} \|SBX - Y\|_2^2 + \sum_{i=1}^{M}\sum_{j=1}^{N} \lambda_{ij} \|\alpha_{ij}\|_0 + \sum_{i=1}^{M}\sum_{j=1}^{N} \beta_{ij} \|X_{ij} - W_{ij}^T.G_{ij}\|_2^2 \qquad (7)$$
$$s.t. \|R_{ij}X - D\alpha_{ij}\|_2^2 \le \varepsilon_{ij},$$

Where λ_{ij} and β_{ij} are the constants for all patches which represents as a regularization parameter.

3.1 Single Dictionary Training

Figure 1. shows the collection of images used for training dictionary.

Figure 1. Set of training images

Figure 2. shows the flow of single dictionary training diagram using multiple generic images.

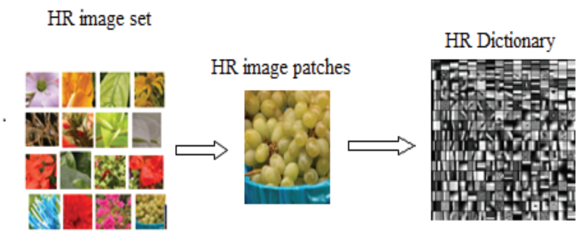


Figure 2. Single dictionary training using multiple generic images

Usually, dictionary is learned from the collection of training images $X = \{x_1, x_2, \ldots\ldots\ldots\ldots x_t\}$. Sometimes, it is very difficult to learn a compact dictionary.

Joint dictionary training is extremely time consuming. In order to solve this problem, we learn single dictionary training. In this training, dictionary is trained off-line from high-resolution patches random sampled from training images.

Procedure for dictionary training is formulated as follows:

$$D, \{\alpha_k\} = \underset{D, \{\alpha_k\}}{\arg\min} \sum_k \|P_k - D\alpha_k\|_2^2 \qquad (8)$$
$$s.t. \|\alpha_k\|_0 \le L \forall k,$$

Where k is an index number of sampled patches, α_k denotes the sparse representation coefficient for k^{th} patch and L represents the maximal number of non-zero entity. After successfully training dictionary, equation (8) degraded into:

$$X, \{\alpha_{ij}\} = \underset{X, \{\alpha_{ij}\}}{\arg\min} \|SBX - Y\|_2^2 + \sum_{i=1}^{M}\sum_{j=1}^{N} \lambda_{ij} \|\alpha_{ij}\|_0 + \sum_{i=1}^{M}\sum_{j=1}^{N} \beta_{ij} \|X_{ij} - W_{ij}^T.G_{ij}\|_2^2 \qquad (9)$$
$$s.t. \|R_{ij}X - D\alpha_{ij}\|_2^2 \le \varepsilon_{ij},$$

Figure 3. shows the block diagram of proposed work.

3.2 ALGORITHM

1: Given input low-resolution image Y. Call only high-resolution dictionary D.

2: By using Bicubic Interpolation method, find out X_0 which can be represents as initial estimation of image X.

3: For image X, we modify equation (9) as:

$$\{\alpha_{ij}\} = \underset{\{\alpha_{ij}\}}{\arg\min} \sum_{i=1}^{M}\sum_{j=1}^{N} \lambda_{ij} \|\alpha_{ij}\|_0 \qquad (10)$$
$$s.t \|R_{ij}X - D\alpha_{ij}\|_2^2 \le \varepsilon_{ij},$$

This process is called as the sparse coding process.

To solve it, we use dynamic group sparsity which has a better improvable over conventional sparse coding.

4: Again for updated coefficient, equation (9) can be modifying as:

$$X = \underset{X}{\arg\min} \|SBX - Y\|_2^2 + \sum_{i=1}^{M}\sum_{j=1}^{N} \beta_{ij} \|X_{ij} - W_{ij}^T.G_{ij}\|_2^2 \qquad (11)$$
$$s.t \|R_{ij}X - D\alpha_{ij}\|_2^2 \le \varepsilon_{ij},$$

In the right hand side, second term can be converted in matrix form,

$$\beta_{ij} \|(I - W)X\|_2^2, \qquad (12)$$

Where

$$W(r,s) = \begin{cases} w_{ij}^{pq}, if(p,q) \text{ is an element of } G_{ij} \\ 0 \quad \text{else} \end{cases} \qquad (13)$$

Here, r, s are the coordinates of (p, q) and (i, j). For next third term, we convert it into global form which term as similarity item.

$$\|X - Z\|_2^2 \le \varepsilon, \qquad (14)$$

Where Z is the reconstructed image derived from only representation coefficient α_{ij} as:

$$Z = \underset{z}{\arg\min} \sum_{i,j} \|R_{ij}Z - D\alpha_{ij}\|_2^2, \qquad (15)$$

This problem has a least-squares solution which can be given by,

$$Z = \left[\sum_{ij} R_{ij}^T R_{ij} \right]^{-1} \sum_{ij} R_{ij}^T D\alpha_{ij}. \quad (16)$$

After calculating W and Z, we again modify equation (11) as:

$$X = \arg\min_X \|SBX - y\|_2^2 + \tau \|(I - W)X\|_2^2 + \zeta \|X - Z\|_2^2, \quad (17)$$

Where τ and ζ are regularization factor to control the influence of the three parts. But equation (17) is not convex in Z. By using Gradient Descent approach, we can easy to minimize this optimization problem to generate high-resolution image X.

5: Go to Step 3 until coverage.

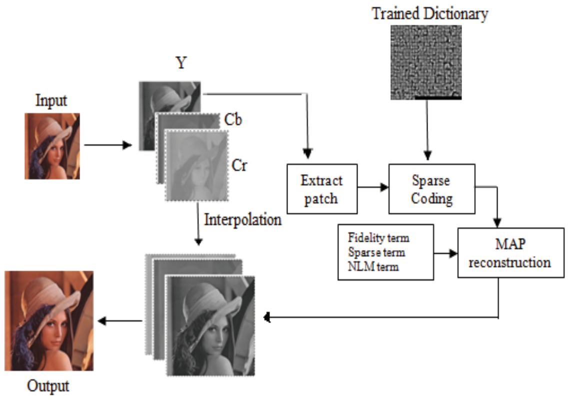


Figure 3. Flow of proposed work

4. Experimental Results And Analysis

In this paper, 2 X magnifications are carried out for input low-resolution images. After completion of these experiments, it is concluded that 5x5 is best patch size. Therefore, we demonstrate experiment on 5x5 patch size. For color images, human eyes are more sensitive to the luminance channel. Hence, we applied proposed algorithm to the luminance channel only.

Here, we trained only single dictionary for high-resolution image patches randomly sampled at rate of 1,00,000 from 69 high-resolution images in the training set. To estimate the performance and computation, we fix dictionary of size 1024 and sparse regularization factor (λ) =0.15 throughout the experiment.

To evaluate the performance of different methods, we use performance parameter such as Peak signal to noise ratio (PSNR), Structural similarity (SSIM) and Image quality index (IQI).

We use Intel(R) Core(TM) i5-3470 CPU@3.20GHz machine with 4GB RAM for simulation.

Table.1. PSNR Performance for Different Methods (Lena)

Upscale	PSNR(dB) Lena 5*5		
	Zoom Image	Bicubic Interpolation	Sparse Recovery
2	30.98	32.79	34.29
4	30.49	32.92	33.51
6	30.06	32.35	32.60

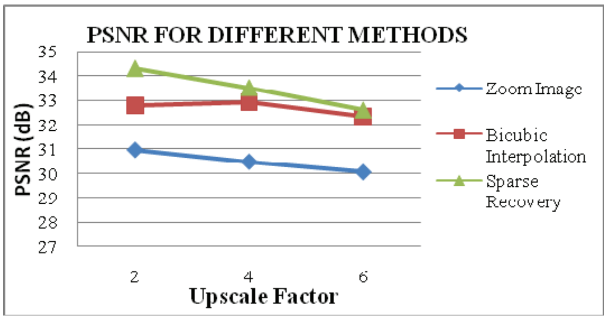


Figure 4. Zoom image, Bicubic Interpolation, Sparse Recovery for Lena image.

Table.1 and Figure 4, indicates the PSNR performance of different methods for 2X, 4X and 6 X magnifications for a Lena image. From table.1, it is illustrated that compare to other methods, sparse recovery is best method which gives superior results.

Table.2. PSNR Performance for Different Methods (Face)

Upscale	PSNR(dB) Face 5*5		
	Zoom Image	Bicubic Interpolation	Sparse Recovery
2	32.22	32.96	34.12
4	32.16	32.72	34.51
6	32.03	33.74	34.17

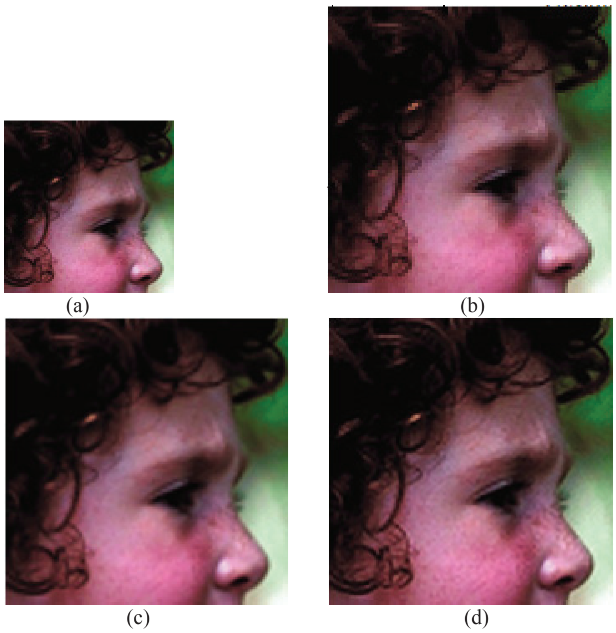


Figure 5. Results of the Face with 2X magnification.
(a) low-resolution (LR) input image,
(b) Zoom Image (PSNR = 32.22dB),
(c) Bicubic Interpolation (PSNR = 32.96dB),
(d) Sparse Recovery for 5x5 patch (PSNR = 34.12dB).

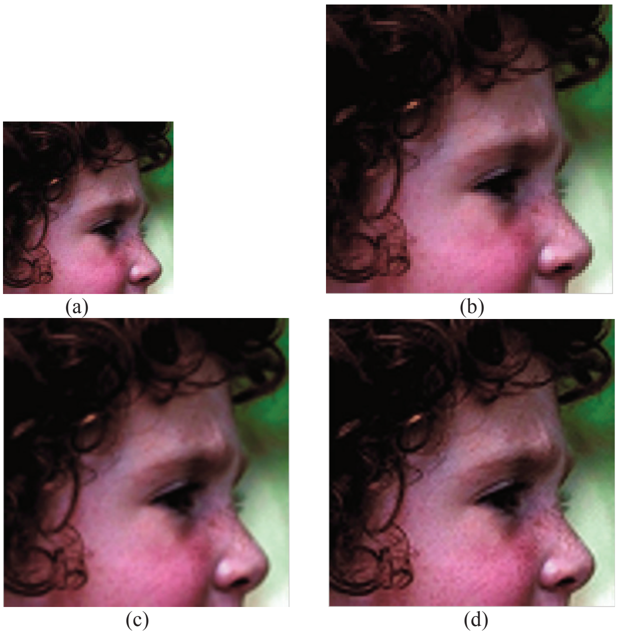


Figure 6. Results of the Face with 4X magnification.
(a) low-resolution (LR) input image,
(b) Zoom Image (PSNR = 32.16dB),
(c) Bicubic Interpolation (PSNR = 32.72dB),
(d) Sparse Recovery for 5x5 patch (PSNR = 34.51dB).

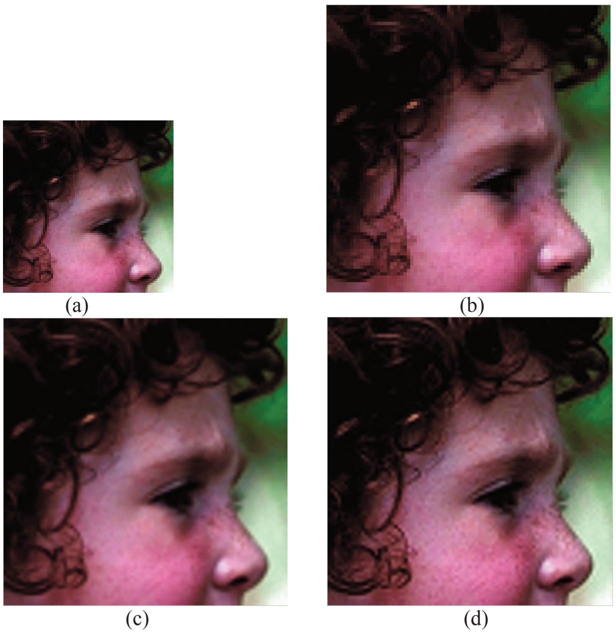


Figure 7. Results of the Face with 6X magnification.
(a) low-resolution (LR) input image,
(b) Zoom Image (PSNR = 32.03dB),
(c) Bicubic Interpolation (PSNR = 33.74dB),
(d) Sparse Recovery for 5x5 patch (PSNR = 34.17dB).

Table 2 and Figures 5, 6, 7, shows more demonstration results of Face image of different methods with 2X, 4X and 6X magnification factor. From figure 5, 6 and 7, it is cleared that sparse recovery is best method which gives best quality results.

From the previous analysis, it is clear that sparse recovery is the best among the other. In this algorithm also we consider 5x5, 7x7 and 9x9 strips. Figures 8, 9 and 10, shows the sparse recovery results along with zoom image and bicubic interpolation method. It is clear from the Figures 8, 9 and 10, that sparse recovery for 9x9 strips is the best. But the run-time is very high.

Figure 8. Results of the Lena with 2X magnification.
(a) low-resolution (LR) input image,
(b) Zoom Image (PSNR = 30.98dB),
(c) Bicubic Interpolation (PSNR = 32.79dB),
(d) Sparse Recovery for 5x5 patch (PSNR = 34.29dB),
(e) Sparse Recovery for 7x7 patch (PSNR = 34.38dB),
(f) Sparse Recovery for 9x9 patch (PSNR = 34.39dB).

Tables 3, 4 and 5, shows more demonstration results of different quality and quantity parameter such as PSNR, SSIM and IQI values for 5x5, 7x7 and 9x9 patch sizes for different test images. From tables 3, 4 and 5, it is cleared that as patch size increases, PSNR value increases little bit but estimated run time also increases. But by considering run time as a main controlling factor with acceptable image quality, 5x5 patch results are more dominating compare to other.

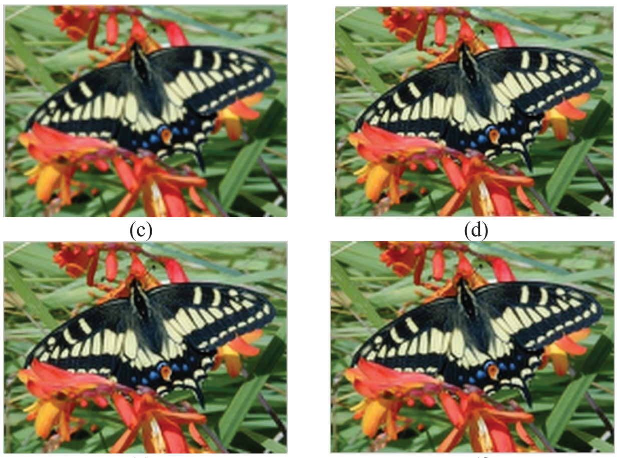


Figure 9. Results of the Butterfly with 2X magnification.
(a) low-resolution (LR) input image,
(b) Zoom Image (PSNR = 28.30dB),
(c) Bicubic Interpolation (PSNR = 33.15dB),
(d) Sparse Recovery for 5x5 patch (PSNR = 36.59dB),
(e) Sparse Recovery for 7x7 patch (PSNR = 36.60dB),
(f) Sparse Recovery for 9x9 patch (PSNR = 36.45dB).

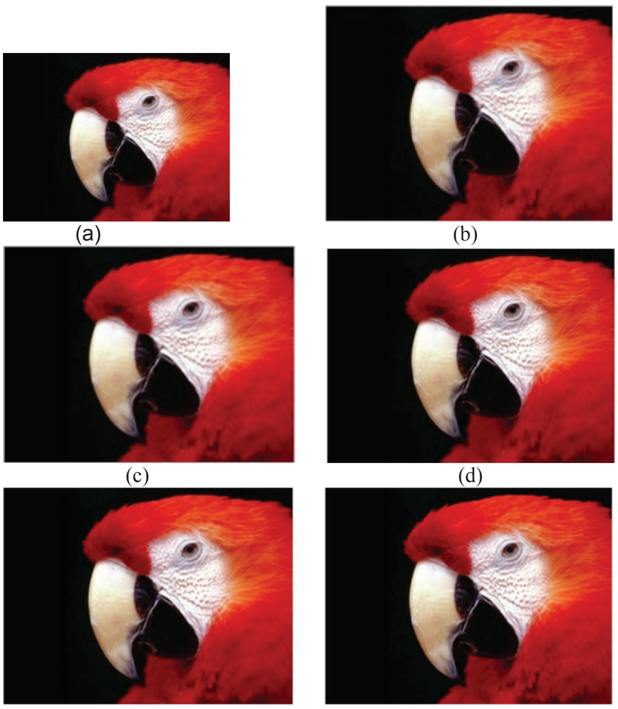


Figure 10. Results of the Parrot with 2X magnification.
(a) low-resolution (LR) input image,
(b) Zoom Image (PSNR = 32.94dB),
(c) Bicubic Interpolation (PSNR = 34.92dB),
(d) Sparse Recovery for 5x5 patch (PSNR = 35.87dB),
(e) Sparse Recovery for 7x7 patch (PSNR = 35.98dB),
(f) Sparse Recovery for 9x9 patch (PSNR = 36.00dB).

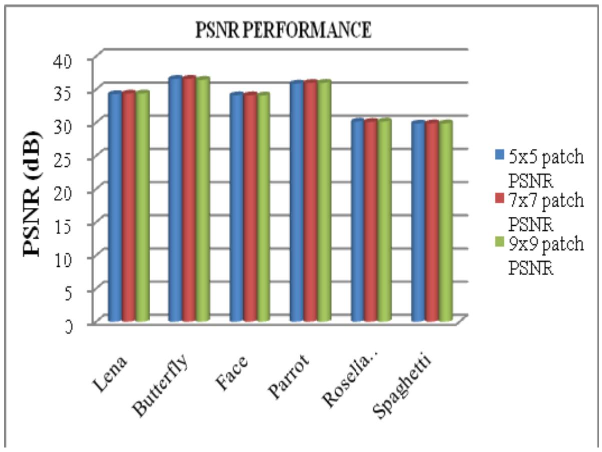


Figure 11. PSNR Performance for patch size 5x5, 7x7 and 9x9 for different test images.

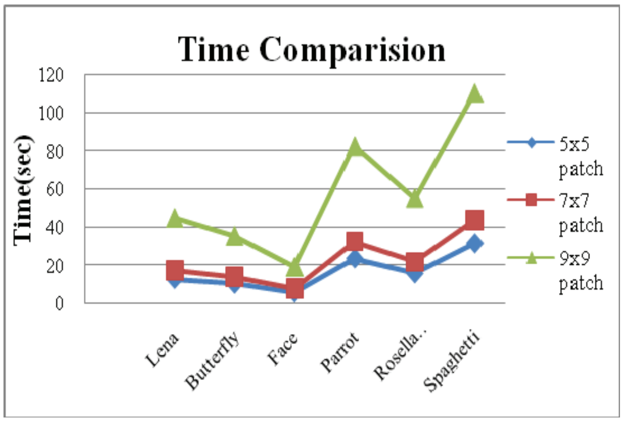


Figure 12. Time Estimation for patch size 5x5, 7x7 and 9x9 for different test images.

Figure 11 and 12, shows the graphical representation of PSNR and run-time. Hence, we consider optimum patch size is 5x5 which provides best super-resolution results and take efficient time.

5. Conclusion

In this paper, a simple unified framework for single image super-resolution using single dictionary is developed. In single dictionary training, dictionary is trained for high-resolution image patches instead of high/low-resolution image patch pairs. Advantage of single dictionary training is no need to train new dictionary again when up-scaling is changed. Another advantage is it reduces the run-time. This dictionary is also called as high-resolution dictionary. From table no. 3, 4 and 5, it is clear that 5x5 patch is the best with acceptable image quality. Using this algorithm, super-resolution results are having rich quality and precise edges. From the results, it is found that single dictionary has done this task more accurately.

Table.3. Performances of Different Methods for Different Test Images for Patch Size 5x5

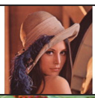
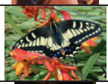
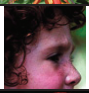
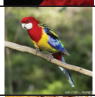
Test images		Zoom method			Bicubic Interpolation			Sparse recovery			Time(sec)
		PSNR	SSIM	IQI	PSNR	SSIM	IQI	PSNR	SSIM	IQI	
Lena		30.98	0.89	0.79	32.79	0.91	0.82	34.29	0.93	0.85	12.876736
Butterfly		28.3	0.92	0.93	33.15	0.97	0.97	36.59	0.98	0.98	10.485493
Face		32.22	0.84	0.75	32.96	0.85	0.76	34.12	0.88	0.80	5.916809
Parrot		32.94	0.92	0.77	34.92	0.94	0.8	35.87	0.95	0.81	23.553663
Rosella-bird		28.55	0.9	0.86	29.58	0.92	0.89	30.16	0.93	0.89	15.802999
Spaghetti		26.87	0.98	0.81	28.95	0.98	0.84	29.85	0.98	0.87	31.365603

Table.4. Performances of Different Methods for Different Test Images for Patch Size 7x7

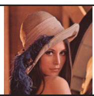
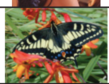
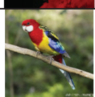
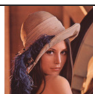
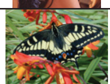
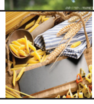
Test images		Zoom method			Bicubic Interpolation			Sparse recovery			Time(sec)
		PSNR	SSIM	IQI	PSNR	SSIM	IQI	PSNR	SSIM	IQI	
Lena		30.98	0.89	0.79	32.79	0.91	0.82	34.38	0.93	0.85	17.639275
Butterfly		28.3	0.92	0.93	33.15	0.97	0.97	36.6	0.98	0.98	14.358423
Face		32.22	0.84	0.75	32.96	0.85	0.76	34.12	0.88	0.8	7.966472
Parrot		32.94	0.92	0.77	34.92	0.94	0.8	35.98	0.95	0.81	32.601400
Rosella-bird		28.55	0.9	0.86	29.58	0.92	0.89	30.11	0.93	0.89	22.012821
Spaghetti		26.87	0.98	0.81	28.95	0.98	0.84	29.89	0.98	0.87	43.989943

Table.5. Performances of Different Methods for Different Test Images for Patch Size 9x9

Test images		Zoom method			Bicubic Interpolation			Sparse recovery			Time(sec)
		PSNR	SSIM	IQI	PSNR	SSIM	IQI	PSNR	SSIM	IQI	
Lena		30.98	0.89	0.79	32.79	0.91	0.82	34.39	0.93	0.85	44.819674
Butterfly		28.3	0.92	0.93	33.15	0.97	0.97	36.45	0.98	0.98	35.484671
Face		32.22	0.84	0.75	32.96	0.85	0.76	34.08	0.88	0.80	19.118581
Parrot		32.94	0.92	0.77	34.92	0.94	0.8	36	0.95	0.81	82.564971
Rosella-bird		28.55	0.9	0.86	29.58	0.92	0.89	30.17	0.93	0.89	55.361026
Spaghetti		26.87	0.98	0.81	28.95	0.98	0.84	29.88	0.98	0.87	110.617177

6. References

[1] Xiaoxuan Chen and Chun Qi, “Nonlinear Neighbor Embedding For Single Image Super-Resolution Via Kernel Mapping”, Signal Process…, no. 94, pp. 6–12, 2014.

[2] Xiaoxuan Chen and Chun Qi, “Low-Rank Neighbor Embedding For Single Image Super-Resolution”, IEEE Signal Process. Letters, vol. 21, no. 1, pp. 79–81, 2014.

[3] M. Elad, “Sparse And Redundant Representation Modeling:What Next?”, IEEE Signal Process. Letters, vol. 19, no. 12, pp. 922–928, 2012.

[4] J. Mairal, F. Bach, and J. Ponce, “Task-Driven Dictionary Learning”, IEEE Trans. Patt. Anal. Mach. Int., vol. 34, no. 4, pp. 791–804, 2012.

[5] M. E. Tipping and C. M. Bishop, “Bayesian Image Super-Resolution”, in Advances in Neural Information and Processing Systems 16 (NIPS), 2003.

[6] R. C. Hardie, K. J. Barnard, and E. A. Armstrong, “Joint Map Registration And High-Resolution Image Estimation Using A Sequence Of Undersampled Images”, IEEE Transactions on Image Processing, vol. 6, pp. 1621– 1633, 1997.

[7] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, “Fast And Robust Multiframe Super-Resolution”, IEEE Transactions on Image Processing, vol. 13, pp. 1327–1344, 2004.

[8] X. Gao, K. Zhang, D. Tao, and X. Li, “Image Super-Resolution With Sparse Neighbor Embedding”, IEEE Trans. Image Process., vol. 21, no. 7, pp. 3194–3205, Jul. 2012.

[9] S. Dai, M. Han, W. Xu, Y.Wu, and Y. Gong, “Soft Edge Smoothness Prior For Alpha Channel Super Resolution”, in Proc. IEEE Conf. Comput. Vis. Pattern Classification, Minneapolis,MN, USA, Jun. 2007, pp. 1–8.

[10] J. Sun, Z. Xu, and H. Shum, “Image Super-Resolution Using Gradient Profile Prior”, in Proc. IEEE Conf. Comput. Vis. Pattern Recog., Anchorage, AK, USA, Jun. 2008, pp. 1–8.

[11] W. T. Freeman, T. R. Jones, and E. C. Pasztor, “Example-Based Super-Resolution”, IEEE comput. Graph. Appl., 22(2): 56–65, 2002.

[12] J. C. Yang, J. Wright, T. Huang, and Y. Ma, “Image Super-Resolution via Sparse Representation”, IEEE. Trans. Image Process.., 19(11): 2861–2873, 2010.

[13] Q. Ning, K. Chen, L. Yi, C. C. Fan, Y. Lu, and J. T. Wen, “Image Super-Resolution Via Analysis Sparse Prior”, IEEE Signal Processing Letters, 20(4): 399–402, 2013.

[14] M. Elad and A. Feuer, “Restoration of A Single Super Resolution Image From Several Blurred, Noisy, And Undersampled Measured Images”, IEEE Trans. Image Process.., 6(12): 1646–1658, 1997.

[15] H. A. Aly and E. Dubois, “Image Up-Sampling Using Total-Variation Regularization with a New Observation Model”, IEEE Trans. Image Process.., 14(10): 1647–1659, 2005.

[16] X. L. Zhang, K. M. Lam, and L. S. Shen. “Image Magnification Based On A Blockwise Adaptive Markov Random Field Model”. Image and Vision Computing, 26(9): 1277–1284, 2008.

[17] M. Irani, S. Peleg. “Improving Resolution by Image Registration”, CVGIP: Graphical Models and Image Processing, 53(3), pp. 23-239, 1991.

[18] R. R. Schultz, R. L. Stevenson. “A Bayesian Approach To Image Expansion For Improved Definition”, IEEE Trans. on Image Processing, 3(2), pp. 233-242, 1994.

[19] H. Stark, P. Oskoui. “High Resolution Image Recovery From Image-Plane Arrays Using Convex Projections”, Optical Society of America, 11(6), pp. 1715-1726, 1989.

[20] M. Elad and M. Aharon, “Image Denoising Via Sparse And Redundant Representations Over Learned Dictionaries”, IEEE Transactions on Image Processing, vol. 15, pp. 3736–3745, 2006.

[21] Jia-Bin Huang and Ming-Hsuan Yang. “Fast Sparse Representation with Prototypes,” Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, June, 2010, pp.3618-3625.

[22] M. Elad, MAT. Figueiredo, and Y. Ma, “On the Role of Sparse and Redundant Representations in Image Processing,” Proceedings of IEEE, Special Issue on Applications of Compressive Sensing & Sparse Representation, June 2010.

[23] J. Mairal, M. Elad, and G. Sapiro, “Sparse Representation for Color Image Restoration,” IEEE Trans. on Image Processing, vol. 17, no. I, pages 53-69, Jan. 2008.

[24] J. Mairal, F. Bach, J. Ponce, G. Sapiro and A. Zisserman, “Non-Local Sparse Models for Image Restoration,” International Conference on Computer Vision, Tokyo, Japan, 2009.

[25] M. Aharon, M. Elad, and A. Bruckstein, “K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation,” IEEE Trans. Signal Process., vol. 54, no. II, pp. 4311-4322, Nov. 2006.

[26] J. Mairal, G. Sapiro, and M. Elad, “Learning Multiscale Sparse Representations for Image and Video Restoration,” SIAM Multiscale Modeling and Simulation, vol. 7, no. I, pages 214-241, April 2008.

[27] R. Rubinstein, M. Zibulevsky, and M. Elad, “Double sparsity: Learning Sparse Dictionaries for Sparse Signal Approximation,” IEEE Trans. Signal Processing, vol. 58, no. 3, pp. 1553-1564, March 2010.

[28] J. Mairal, F. Bach, 1. Ponce, G. Sapiro, and A. Zisserman, “Supervised dictionary learning,” Advances in Neural Information Processing Systems (NIPS '08), pp. 1033-1040.

[29] G. Monaci and P. Vanderqheynst, “Learning structured dictionaries for image representation,” in Proc. IEEE Int. conf Image Process. pp. 2351- 2354, Oct. 2004.

[30] R. Rubinstein, A.M. Bruckstein, and M. Elad, “Dictionaries for sparse representation modeling,” Proceedings of IEEE, Special Issue on Applications of Compressive Sensing & Sparse Representation, vol. 98, no. 6, pp. 1045-1057, June, 2010.

[31] G. Mu, X. Gao, K. Zhang, X. Li, D. Tao., “Single Image Super Resolution with high resolution dictionary,” in Proc. IEEE Int.conf Image Process.pp.1141-1144, Sept. 2011.

[32] Buades, A., Coll, B., and Morel, J.M., “A non-local algorithm for image denoising,” In CVPR, pages 60-65, Jun. 2005.

Biographies

Mr. Kiran Jadhav has received Bachelor Degree in Electronic and Telecommunication Engineering from Pune University. He is Master Degree student at University of Mumbai in the Electronic and Telecommunication Engineering. His research interest in Image Processing.

Dr. Ramesh Kulkarni is a PhD scholar in the Department of Electronics & Communication Engineering, National Institute of Technology, Rourkela, Orissa, India. He received his bachelor degree in Electronics & Communication from Mysore University and master degree in Digital Electronics form Karnataka University, Karnataka. He has many publications in international journals and international conferences. His research area is Image processing, Non – Linear filters, and Digital signal processing. He is presently working with VES Institute of Technology, Mumbai as Professor in the Electronics & Telecommunication Department.

Mr. Gaurav Tawde has received bachelor degree in Electronics & Telecommunication Engineering from Pune University and Master degree in Electronics & Telecommunication Engineering form Mumbai University, India. His areas of interest are Digital Electronics, Microcontrollers and Image Processing. He is a life member of ISTE and IETE. He is presently working with VES Institute of Technology, Mumbai as Assistant Professor in the Electronics & Telecommunication Department.

Graphics, Vision and Image Processing, V. 15, No. 2, ISSN 1687-398X, Delaware, USA, December 2015

www.icgst.com

Supervised Spectral-Spatial Hyperspectral Image Classification based on Oversampling and Composite Kernels

Rafika Ben Salem[1], Karim Saheb Ettabaa[2,3] and Mohamed Ali Hamdi[1]

[1]*Laboratory for Materials, Molecular and applications, National Institute of Applied Sciences and Technology, Tunis, Tunisia.*

bs_rafika@yahoo.fr, mohamedalihamdi@yahoo.fr

[2]*Laboratory for research in computer Arabized and integrated documentation, National School of Computer Sciences, Tunis, Tunisia.*

[3]*Laboratory ITI, Telecom Bretagne, Brest Iroise Technopole CS 81828*

karim.sahebettabaa@riadi.rnu.tn

Abstract

The limited availability of labeled training samples for supervised hyperspectral image classification decrease the accuracy of supervised classifier. To overcome this difficulty, this paper introduces a new supervised classification approach for remotely sensed hyperspectral image data which process accurately with limited number of training samples and integrates the spectral and spatial information via composite kernel. In fact, the developed method introduces, in the learning step, new examples oversampled from the available limited training set by using interpolation techniques. First, each pixel will be presented by two vectors: spectral vector containing all the spectral information and spatial vector including contextual information extracted using Extended Multi-attribute profiles (EMAP). Then, an interpolation technique is used to generate new training samples, using the limited available data. Finally, a support vector machines (SVMs) with composite kernel is efficiently trained to generate the classification map. The proposed classification approach is evaluated using both simulated and real hyperspectral data sets, allowing higher performance when compared with the classification without oversampling. The integration of interpolation methods with SVMs, combined with the use of spectral and spatial information, represents a contribution in the literature. This approach is shown to provide accurate classification of hyperspectral imagery with limited number of training samples.

Keywords: *Hyperspectral images, SVM, composite kernels, interpolation techniques.*

Nomenclature:

AA	Average Accuracy
AP	Attribute Profile
CFS	Correlation-based Feature Selection
EAP	Extended Attribute Profile
EMAP	Extended Multi-Attribute Profile
ICA	Independent Components Analyses
k	Kappa coefficient
Lp	Linear Projection
mRMR	Minimum-Redundancy-Maximum-Relevance
OA	Overall Accuracy
PCA	Principal Components Analyses
SVM	
SVM-RFE	Support Vector Machine Recursive Feature Elimination
$C=\{1,...,C\}$	set of C classes
$P=\{1,...,n\}$	Set of integer indexing the n pixels of an image.
$Y = (y_1,...,y_n)$	Image with n pixels
$Y^{spect} = (y_1^{spect},...,y_n^{spect}) \in R^{d\times n}$	Spectral features: each pixel is characterized by a d-dimensional spectral vector.
$Y^{spat} = (y_1^{spat},...,y_n^{spat}) \in R^{m\times n}$	Spatial features: each pixel is characterized by a m-dimensional spatial vector.
$Lab = (lab_1,...,lab_n)$	Labels of n pixels.
$T_i^{spect} = \{(i, y_{i\,1}^{spect}),...,(i, y_{i\,t}^{spect})\}$	Set of t spectral vectors of available labeled samples in class i.
$T_i^{spat} = \{(i, y_{i\,1}^{spat}),...,(i, y_{i\,t}^{spat})\}$	Set of t spatial vectors of available labeled samples in class i.

$X_i^{spect} = (x_{i\,1}^{spect},\ldots,x_{i\,t}^{spect}) \in R^{d\times t}$	Set of real presenting the abscissas of t spectral training samples in class i.
$X_i^{spat} = (x_{i\,1}^{spat},\ldots,x_{i\,t}^{spat}) \in R^{m\times t}$	Set of real presenting the abscissas of t spatial training samples in class i.
$Y_{i\,new}^{spect} = (y_1^{spect},\ldots,y_g^{spect}) \in R^{d\times g}$	g generated spectral vectors for class i used to train the spectral kernel.
$Y_{i\,new}^{spat} = (y_1^{spat},\ldots,y_g^{spat}) \in R^{m\times g}$	g generated spatial vectors in class i required for the learning of the spatial kernel.
K^{spect}	spectral kernel
K^{spat}	spatial kernel
f	Interpolation function.
μ	Weight $0<\mu<1$

1. Introduction

Supervised classification of high dimensionality remotely sensed hyperspectral images that requires prior knowledge is a challenging task [1]. Hughes phenomenon manifests as the data dimensionality increases. This problem refers to the unbalance between the high number of available spectral bands and the limited availability of labeled training samples used for the learning step of the classifier. In order to overcome this problem, two strategies have been conducted: either by reducing the dimensionality of images or by increasing the number of training samples. For the first strategy, several features selection [2] and extraction methods have been combined with machine learning techniques including SVMs. In [3], a significant increase of the classification accuracy has been resulted according to the combination of SVMs with four features selection methods which are SVM-RFE, CFS, mRMR and Random Forest. ICA has been investigated in [4] to reduce the dimensionality of the hyperspectral image. In [5], the non supervised features selection method Lp has been combined with SVM to improve the performance of the classification. For the second strategy, various semisupervised approaches have been presented in the literature. Indeed, the increase of the number of training examples presents a challenge because the collection of labeled data is generally difficult, expensive and time-consuming. For that, semi-supervised learning techniques [6] have been adopted in hyperspectral image classification. These approaches are based on the exploitation of both training data and unlabeled patterns. In [7], active queries have been used in the development of a semiautomatic procedure to generate land cover maps from remote sensing images. A significant improvement has been achieved according to the exploitation of active learning algorithms in the development of semi-supervised self-learning approach presented in [8].

In this work, we propose another solution to have accurate classification with small training set by increasing the number of labeled samples without using semi-supervised learning techniques. The main idea of the proposed method is the oversampling of the available learning samples to generate new training data. In this context, interpolation techniques can provide competitive advantages by creating new samples over curve that interpolates the existing labeled examples. This idea is not exploited yet in the literature to perform the accuracy of the supervised hyperspectral image classification.

Combining spectral and spatial information is a recent trend that proved successfully results in several recent remote sensing hyperspectral images classification studies. Hence, many spectral-spatial classification approach have been presented such as composite kernels presented in [9], the generalized composite kernel framework proposed in [10], the SVM ensemble approach combining spectral, structural, and semantic features developed in [11], the introduce of spectral and different spatial features via the composition of kernels in [12] and the investigation of segmentation methods in [13]. In particular, composite kernels have shown high performance in term of accuracy and computational time.

In this paper, we propose a new supervised hyperspectral image classification approach which overcomes the problem of the limited number of training samples by generating new training samples and combines spectral and spatial information via composite kernels. The algorithm implements the following three main steps: 1) spectral and spatial characterization, where each pixel is presented by spectral and spatial vectors, 2) oversampling which generate new training samples by means of interpolation methods, and 3) classification that results the classification map built by SVMs with composite kernels. The main novelty of our proposed work is the integration of interpolation techniques to generate from a small training set an important number of new learning examples, which will be shown to increase the size of labeled patterns and to provide a good classification results.

The remainder of this paper is organized as follows. Section 2 formulates the problem. Section 3 describes the proposed approach. Section 4 reports classification results based on simulated and real hyperspectral data sets. Finally, Section 5 concludes with some remarks.

2. Problem formulation

Supervised classification aims at assigning a label $lab_i \in C$ to each pixel y_i by using a set of labeled samples called training samples required to train the classifier. This process results in an image of class labels *Lab*.

Supervised classifiers SVMs are particularly applied in the remote sensing field due to their ability to successfully manipulate small training data sets [14, 15]. However, in the case of hyperspectral images various researches [1] proved that the internal class variability (various spectral signatures in the same class) and the high dimensionality of the data require an important number of training samples to have accurate classification. In fact, SVM aims to find a hyperplane that separates the dataset into predefined number of classes based on training examples [16]. The optimal separation hyperplane refers to the decision boundary that minimizes misclassifications, obtained in training step. Then, if the number of learning examples is limited, the classifier cannot find the optimal separation

hyperplane producing thus an important rate of misclassifications.
In the other hand, the collection of learning samples involves expensive ground campaign that conduct to have a small set of labeled examples. To overcome this limitation and increase the size of training set, interpolation techniques can be exploited to generate synthetic data using the limited number of available samples. Indeed, new samples can be created under curves that interpolate the t points presenting the set of training samples.
Interpolation process consists to construct a function $f(x)$ that passes through k points: $(x_1, f(x_1)),..., (x_k, f(x_k))$. Various interpolation methods have been proposed in literature such as:

- Linear interpolation: For each interval (x_i, x_{i+1}), the interpolation formula is given as Equation (1):

$$f(x) = Af(x_i) + Bf(x_{i+1}) \tag{1}$$

where

$$A = \frac{x_{i+1} - x}{x_{i+1} - x_i}, \quad B = 1 - A = \frac{x - x_i}{x_{i+1} - x_i} \tag{2}$$

- Lagrange interpolation: In this case, the interpolation polynomial is given as a linear combination of Lagrange basis polynomials $l_j(x)$ (Equation (3)):

$$f(x) = \sum_{j=1}^{k} f(x_j) l_j(x) \tag{3}$$

with

$$l_j(x) = \frac{\prod_{\substack{i=0 \\ i\neq j}}^{k} (x - x_i)}{\prod_{\substack{i=0 \\ i\neq j}}^{k} (x_j - x_i)} \tag{4}$$

- Cubic spline interpolation: A cubic spline is a spline for which the function is a polynomial of degree=3 on every subinterval (Equation (5)).

$$f(x) = \begin{cases} f(x_0), & x \in [x_0, x_1) \\ f(x_1), & x \in [x_1, x_2) \\ \vdots \\ f(x_{k-1}), & x \in [x_{k-1}, x_k) \end{cases} \tag{5}$$

where $\forall$ i, the degree of $f(x_i)$ is equal to 3.

Following the interpolation techniques described earlier, we have developed a new supervised classification method which generates, in the first step, a new set of spectral and spatial samples through a limited number of available labeled data in the goal of increasing the size of learning set indispensable to train the two different kernels K^{spect} and K^{spat} and results, in the second step, the final classification map will be generated according to the use of SVMs with composite kernels. In our proposed algorithm, the spectral information is represented by Y^{spect} which is trained by a spectral kernel. The spatial information is presented by Y^{spat} which is processed by a spatial kernel. In each class i, generated spectral information $Y^{spect}_{i\ new}$ is resulted from the interpolation of $\{(x^{spect}_{i\,1}, y^{spect}_{i\,1}),...,(x^{spect}_{i\,t}, y^{spect}_{i\,t})\}$ and the set of new spatial samples $Y^{spat}_{i\ new}$ is generated from the interpolation of $\{(x^{spat}_{i\,1}, y^{spat}_{i\,1}),...,(x^{spat}_{i\,t}, y^{spat}_{i\,t})\}$.

3. Proposed approach

In this section, we present, first, spectral and spatial characterization techniques used to extract features. Second, we propose new algorithm which uses interpolation methods to generate new learning samples. Finally, we integrate the spectral and the spatial information via SVM and composite kernels.

3.1. Spectral and spatial characterization

The rich spectral and spatial information available in hyperspectral images allows for the possibility to distinguish between spectrally similar materials. Various methods have been widely used in the literature for spectral and spatial characterizing hyperspectral pixels. For the spectral characterization, authors usually used all the spectral information or dimensionality reduction techniques like PCA and ICA to extract the most informative data. For the spatial features extraction, different means have been adopted such as: features provided from the neighborhood of the pixel, morphological filters and attribute filters. In this paper, we focus on the uses of all the spectral information for the spectral characterization and EMAP based on attribute filters for the spatial characterization. EMAP has been selected due to their capability to distinguish between similar objects and to model different spatial structures in the scene.
EMAP [17] is a profile that stacked the EAPs obtained using different type of attributes. The EAP is resulted by generating an AP (obtained by applying a sequence of attribute filters using various thresholds) on each of the first p principal components.

3.2. Generation of synthetic data

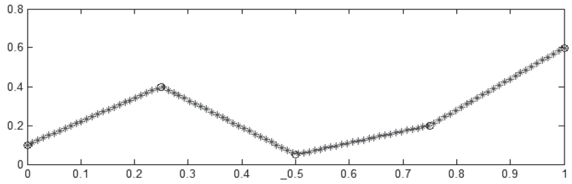
One of the main problems in hyperspectral image classification is the presence of Hughes phenomenon introduced by the unbalance between the high number of available spectral bands and the limited availability of labeled training samples. To overcome this difficulty, we process in this paper to investigate interpolation techniques for generating new training samples from the available limited number of learning examples. Then, new samples can be created under curves that interpolate the t points presenting the set of t training samples

(Figure 1).

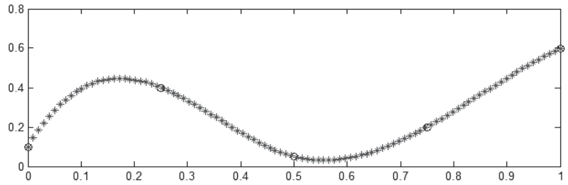
We assume that each training sample i is presented by a f dimensional vector $y_i = (y_i^1, \ldots, y_i^f)$ combining the f features. Hence, generated samples must be a f dimensional vectors $y^{new} = (y^{1new}, \ldots, y^{f\ new})$ that combining the new values of the f features resulted from the interpolation process. For that, each feature j of the training sample i must be presented by a point $P(x^j_i, y^j_i)$. The evaluation of the function f that interpolates training samples points in g abscissas $x^{j\ new}$ generates the new samples y^{jnew}. Based on the aforementioned assumptions, we have the following generative function (Equation (6)):

$$y^{j\ new} = f(x^{j\ new}) \qquad (6)$$

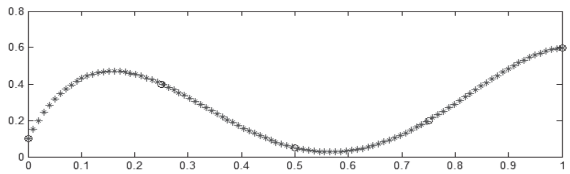
In this paper, we focus on three interpolation techniques: linear interpolation, cubic spline interpolation and Lagrange interpolation. Algorithm 1 shows the pseudocode for the learning data generation based interpolation techniques.



Linear interpolation



Cubic spline interpolation



Lagrange interpolation

Figure 1. Representation of generated data under interpolation curves ('*' illustrate new samples, 'o' illustrate the available data).

Algorithm 1: *Oversampling*

Input: $Y^{train}=(y_1, \ldots, y_t)$: Training samples set.
f: dimension of vectors.
Output: $Y^{new}=(y^{new}_1, \ldots, y^{new}_g)$: Generated training samples.

for i=1 to f **do** (for each feature i)

$y = Y^{train}(i,:)$
$x^{train} = (x_1, \ldots, x_t)$
} Coordinate of the points to be interpolated.

$y=f(x^{train})$ (compute the interpolation function according to (1) or (3) or (5))
$x^{new}=(x^{new}_1, \ldots, x^{new}_g)$: new samples abscissas
$N=f(x^{new})$ (compute new training samples according to the evaluation of the function f in new abscissas)
$Y^{new} = [Y^{new}; N]$
end

3.3. Spectral spatial classification via SVM with composite kernels

According to their high performance registered on the process of data with high dimensionality, SVMs have been widely adopted in the classification of hyperspectral images.

SVM [18] is a kernel based classifier consisting in projecting data in a higher dimension space by means of non-linear mapping function Φ and aiming at finding the optimal separator hyperplan by margin maximization. SVM has been proposed first for binary classification, after it has been introduced to solve multi-class classification.

Given a labeled training data set $\{(x_1, L_1), \ldots, (x_m, L_m)\}$, where $x_i \in R^N$ and $L_i \in \{-1, 1\}$. The SVM classifier consist in calculating the optimum separating hyperplan defined by Equation (7).

$$w\,x + b = 0 \qquad (7)$$

where (w, b) are the parameters of the hyperplan.

Thus, the classifier can be defined as Equation (8):

$$h(x,w,b)=\mathrm{sgn}(w*x+b) \qquad (8)$$

The support vectors lie on two hyperplanes of equation: $w\ x + b = +-1$. The maximization of the margin leads to the following optimization problem (Equation (9)):

$$\min\{\frac{1}{2}\|w\|^2\} \text{ with } L_i\,(w\,x + b) \geq 1,\ i=1....m \qquad (9)$$

If the training samples are not linearly separable, the optimization problem can be solved by using Lagrange multipliers λ_i and becomes (Equation (10)):

$$\begin{cases} \min\left\{\sum_{i=1}^{m}\lambda_i - \frac{1}{2}\sum_{i,j=1}^{m}\lambda_i\lambda_j L_i L_j x_i x_j\right\} \\ 0 \leq \lambda_i \leq C, \forall i = 1,2,\ldots,m \\ \sum_{i=1}^{m}\lambda_i L_i = 0 \end{cases} \qquad (10)$$

with C is a regularization parameter introduced to reduce the weight of misclassified vectors. Thus the classifier function can be computed by optimization process. To solve this difficulty SVM can be generalized to compute nonlinear decision surface by projecting the data in a higher dimension space where they are considered to become linearly separable. The projection is allowed by using non linear function Φ and can be simulated using a kernel method. Hence, the dot products $(x_i.x_j)$ involved in Equation (10) is replaced by $K(x_i.x_j)= \Phi(x_i)\,.\,\Phi(x_j)$.

Then the non linear classifier can be expressed as Equation (11):

$$h(x, w, b) = \text{sgn}\left(\sum_{i=1}^{N_s} \lambda_i y_i K(S_i, x) + b\right) \quad (11)$$

with S_i are the N_s support vectors and K is a kernel that satisfies Mercer's conditions such as linear, polynomial and Radial Basis Function (RBF) kernels.

Some properties of Mercer's kernel: let K_1, K_2 two valid Mercer's kernels and $\alpha>0$, Then $K(x_i.x_j)=K_1(x_i.x_j)+K_2(x_i.x_j)$ and $K(x_i.x_j)= \alpha\ K_1(x_i.x_j)$ are a Mercer's kernels.

In order to improve the classification performance achieved by using the spectral information alone, various spectral-spatial classification approaches incorporating the spatial information in addition to the spectral information have been proposed in the literature. In particular, the uses of SVM with composite kernels that combining spectral kernel K_ω for the spectral features x^w and spatial kernel K_s for the contextual features x^s has shown high performance in term of accuracy and computational time.

Composite kernels [9] take advantage of the direct sum of Hilbert spaces by which two (or more) Hilbert spaces H_k can combine into a larger Hilbert space. This allows proposing tree different composite kernels:

- Direct summation kernel (Equation (12)):

$$K(x_i, x_j) = K_w(x_i^w, x_j^w) + K_s(x_i^s, x_j^s) \quad (12)$$

- Weighted summation kernel (Equation (13)):

$$K(x_i, x_j) = \mu K_w(x_i^w, x_j^w) + (1-\mu) K_s(x_i^s, x_j^s) \quad (13)$$

with $0 < \mu < 1$

- Cross-information kernel (Equation (14))

$$K(x_i, x_j) = K_w(x_i^w, x_j^w) + K_s(x_i^s, x_j^s) + K_{ws}(x_i^w, x_j^s) + K_{sw}(x_i^s, x_j^w) \quad (14)$$

In this case spectral and spatial vector must have the same dimension.

It's notable that composite kernels are Mercer's kernels. For that, they have been used to solve spectral-spatial classification with SVM.

3.4. Supervised classification algorithm

To summarize the description of our proposed method, Algorithm 2 provides a pseudocode for our newly developed spectral spatial supervised classification algorithm based on a SVM classifier with composite kernels and oversampling. Figure 2 shows the flowchart of this algorithm.

Figure 3 illustrates a block diagram summarizing the most relevant steps of the newly proposed classification algorithm.

Algorithm 2 SVM_oversampling

Input: $Y = (y_1, ..., y_n)$, Y^{train}

Output: $L = (l_1, ..., l_n)$

Y^{spect} = Spectral characterization(Y)

T_i^{spect} = Spectral characterization(Y^{train})

Y^{spat} =Spatial characterization(Y)

T_i^{spat} = Spatial characterization(Y^{train})

T^{spat}= T^{spect}= []

for i=1 **to** c **do**

$Y_{i\ new}^{spect}$ =oversampling(T_i^{spect} , n)

$Y_{i\ new}^{spat}$ =oversampling(T_i^{spat} , m)

T_i^{spect} =[T_i^{spect} , $Y_{i\ new}^{spect}$]

T_i^{spat} =[T_i^{spat} , $Y_{i\ new}^{spat}$]

T^{spect}=[T^{spect}, T_i^{spect}]

T^{spat}=[T^{spat}, T_i^{spat}]

end

L= Classification(Y^{spect} , Y^{spat} , T^{spat}, T^{spect})

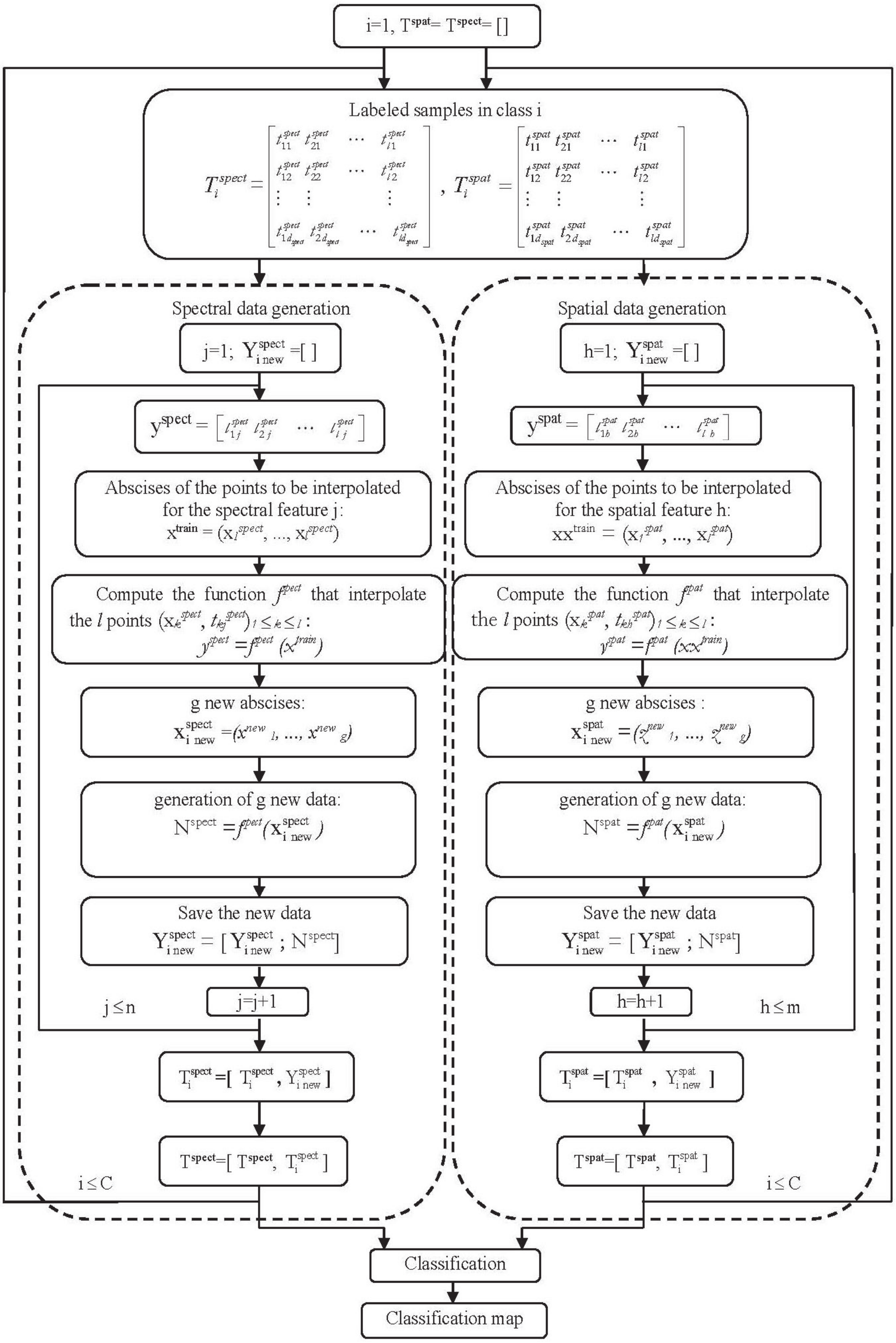


Figure 2: flowchart/block diagram of the proposed classification algorithm

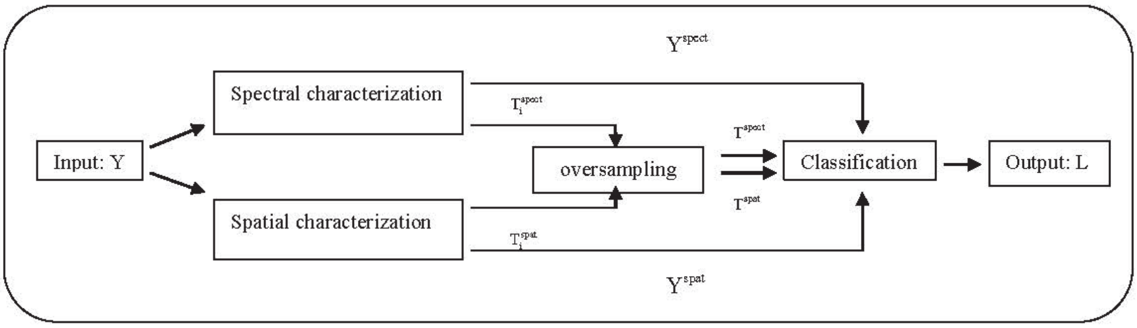


Figure 3. Block diagram summarizing the most relevant steps of the proposed classification algorithm.

4. Experimental results

This section uses both simulated and real hyperspectral data sets to illustrate the effectiveness of the proposed spectral spatial classification algorithm in different analysis scenarios. The remainder of this section is organized as follows. Section 4.1 first explains the parameter settings adopted in our experimental evaluation. Section 4.2 then evaluates the proposed algorithm by using simulated data sets, whereas Section 4.3 evaluates the proposed classification algorithm using real hyperspectral image.

4.1. Parameter Settings

Before presenting our results with simulated and real hyperspectral data sets, we discuss first the parameter settings adopted in our experiments. In all experiments, we used RBF and polynomial kernels for the spectral and spatial features, respectively, to construct composite kernels. The training sets are randomly selected from the available labeled samples and that the remaining samples are used for validation. Ten Monte Carlo runs have been conducted, the OA (in percent)) is obtained after each run. The labeled samples for each Monte Carlo simulation are obtained by resampling the available learning examples. For the SVM parameters and the parameters used for building EMAP, we optimized the SVM parameters using tenfold cross-validation and we used the first three PCs which account for most of the variance present in the considered data sets to compute EMAP.

4.2. Experiments with Simulated Hyperspectral Data

In our experiments, we have used a 30×30 simulated hyperspectral scene containing 4 classes. Adopted spectral signatures are obtained from the U.S. Geological Survey (USGS) digital spectral library (Figure 4). EMAP were built using threshold value in the range of 10% - 70% with a step of 20% for the standard deviation attribute and thresholds of 30, 40, 50 and 60 for the area attribute. In the learning step, we used new data (*g* samples) generated according to the interpolation of *l* training samples in each class. Hence, the total number of training samples is *(l+g)*.

We have conducted four different experiments with the simulated hyperspectral image. These experiments have been conducted to analyze several relevant aspects of our proposed classification method.

1) In our first experiment, we evaluate the impact of the adopted interpolation method on the classification result.

2) In our second experiment, we evaluate the impact of the available training set size on the classification result.

3) In our third experiment, we evaluate the impact of the number of generated training samples on the classification result.

4) In our fourth experiment, we analyze the impact of the parameter μ adopted for weighted summation kernel on the classification output.

In all these experiments, we will use the average and the standard deviation of classification accuracy (OA) and kappa coefficient (Kappa) obtained after ten Monte Carlo runs as a references to evaluate the performance of the proposed classification approach.

$$\text{OA} = \frac{\text{number of pixels correctly classified}}{\text{Total number of pixels}} * 100$$

$$\text{Kappa} = \frac{\text{Number of pixels correctly classified}}{\text{Number of pixels correctly classified} + \text{Number of confusion}} * 100$$

Figure 4. Ground truth of the simulated hyperspectral scene.

4.2.1. Experiment 1—Impact of the applied interpolation method

In this experiment, we analyze the impact of the interpolation techniques used to generate new training data. Two methods must be applied: method to interpolate spectral samples and another to interpolate spatial samples. To address this issue, we analyze the performance of the proposed method for different combination of interpolation methods in a classification problem with generating in each class 100 samples about 50 available training samples and with using direct

summation kernel to combine spectral and spatial features. Figure 5 shows the OAs and kappa obtained by the proposed classification algorithm according to the applied combination. Notice the good performance achieved by the proposed classification algorithm, which yielded much better OA and kappa results (OA= 94,48% and kappa=97,29%) in case with linear interpolation for spectral samples and cubic spline interpolation for spatial examples. Furthermore, all the applied combinations allowed to have accurate results than the classification without oversampling (OA>OA_without= 75,86% and kappa> kappa_without= 85%). This is reasonable since the performance of the classification is increased as the number of training samples increases. This also indicates the advantage of generated training data which increases the capability of SVM to find the optimal separator hyperplan. The robustness of the proposed methods in the presence of very limited training sets is analyzed in more detail in the following experiment.

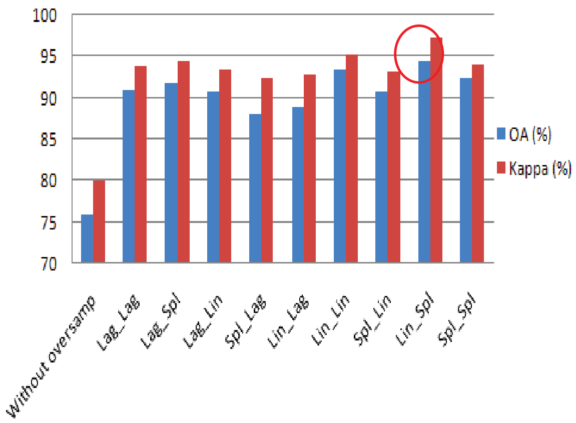
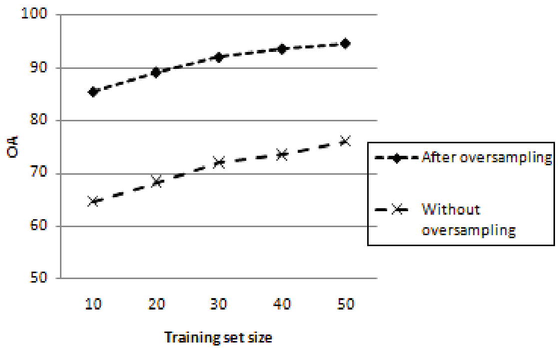


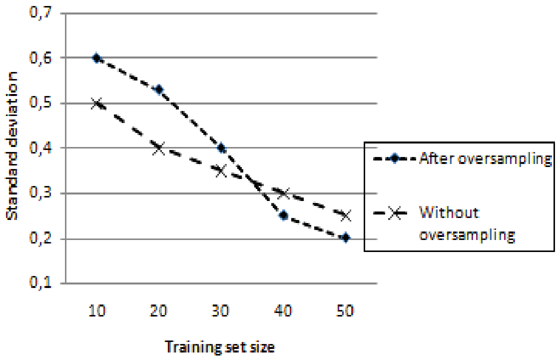
Figure 5. Variation of OA and Kappa coefficient according to the variation of the adopted interpolation techniques.

4.2.2. Experiment 2—Impact of the available training set size

In our second simulated image experiment, we analyze the impact of the training set size on the classification performance. Figure 6 (a) and (b) shows the OA and standard deviation (Std) results, respectively, obtained by our proposed methods as a function of the number of labeled samples (*l*) used in the training process with generating 100 samples about *l* according to use Lin-Spl combination and with using direct summation kernel. Notice the quality of the classification results obtained by our proposed algorithm, which shows high robustness even with very limited training set sizes. As the number of labeled samples increases, the OA increases and the standard deviation decreases. This is expected since an increase of the number of labeled samples should decrease in the uncertainty when estimating the optimal separator hyperplan.



(a)



(b)

Figure 6. Variation of OA (a) and standard deviation (b) with the increase in the size of the available training set.

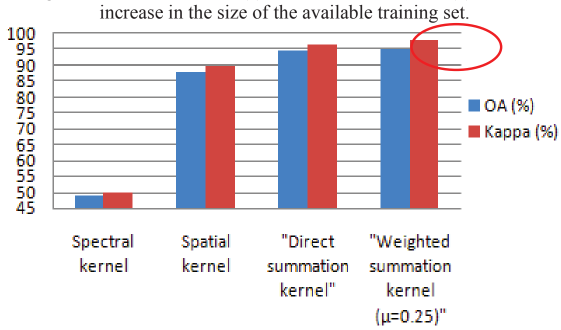
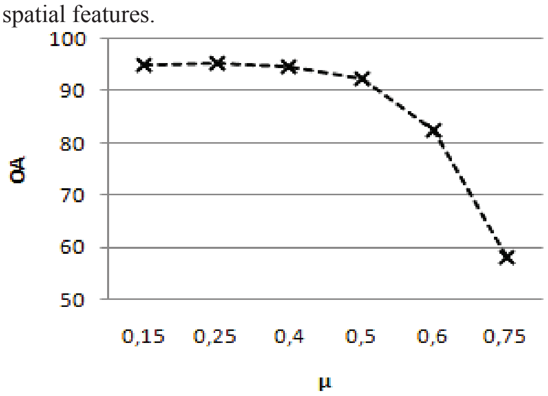


Figure 7. Resulted OA and Kappa coefficient according to the adopted kernel type

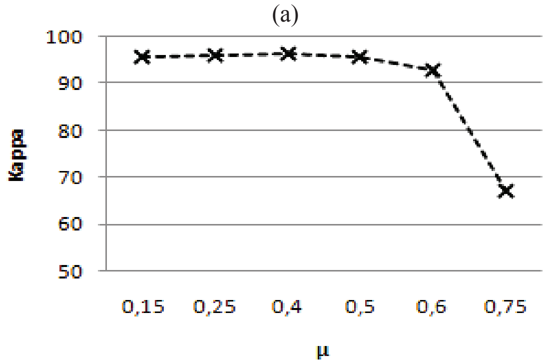
4.2.3. Experiment 3—Impact of the adopted composite kernel type

In our third simulated image experiment, we analyze the impact of the composite kernel type on the classification performance. For that, we applied two different composite kernels which are direct summation kernel and weighted summation kernel. The cross-information kernel can't be applied because the spectral and the spatial vectors have different dimension. Figure 6 shows the OA and kappa coefficient results, respectively, obtained by our proposed methods as a function of adopted kernel type. As shown in Figure 7, the performance of the proposed classification algorithm increases when using composite kernels which combine spectral and spatial features. Furthermore, we can note that the weighted summation kernel introducing a trade-off (µ) between spectral and spatial kernels with µ=0.25 performs more accurately than the direct summation kernel.

To address the impact of the parameter μ intended to perform the spectral information against the spatial information, we analyze the performance of the proposed classification method for different values of μ. Figure 8 (a) and (b) shows the OA and kappa coefficient results, respectively, obtained by our proposed methods as a function of μ. Notice the good performance achieved by the proposed classification algorithm, which yielded much better OA an kappa results (OA= 95,34% and kappa=96,6%) in cases with μ=0.25. This indicates that the proposed method performs accurately when we valorize the spatial information against the spectral information which illustrates the importance of the spatial features.



(a)

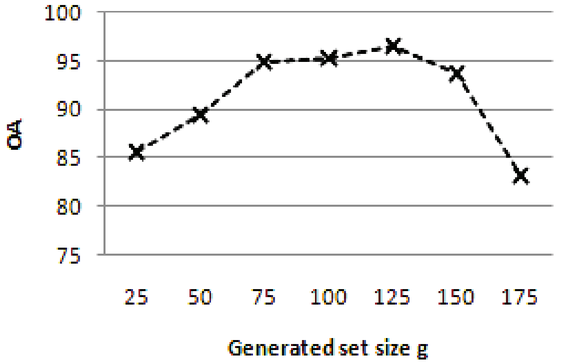


(b)

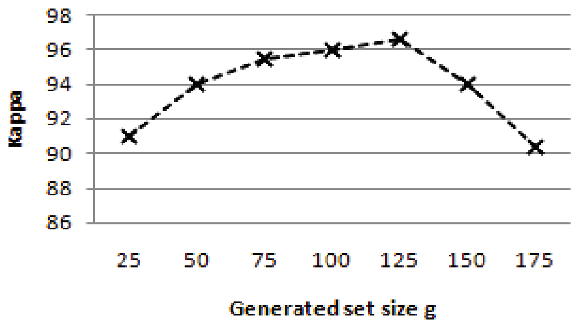
Figure 8. Variation of OA (a) and kappa coefficient (b) according to the variation of μ

4.2.4. Experiment 4—Impact of the generated training samples set size g

In our last experiment with simulated data, we conduct an evaluation of the impact of generated data set size on the proposed classification algorithm by using only *l = 20* labeled samples per class and weighted summation kernel with μ=0,25. Figure 9 (a) and (b) shows the OA and the kappa results as a function of generated data set size. From figure 8, we can conclude that the classification performance indeed depends on the size of generated training samples. With a size 75 ≤ g ≤ 150, the proposed classification method leads to good results (93%<OA<96,6%). This indicates that the method leads to high values of the OA for limited number of training samples (*l=20*). On the other hand, we have experimentally observed that the OA and the kappa results converge to low values for g>150. This indicates that the augmentation of g can leads to poor results lower than those obtained in the case of classification with *l=20,* i.e., without oversampling (OA= 83,5%, kappa=91%). This is because, the increase in the number of generated data from a limited set of available training samples leads to have bad training samples that confuse the classifier.



(a)



(b)

Figure 9. Variation of OA (a) and kappa coefficient (b) according to the increase in generated set size.

4.3. Experiments With real Hyperspectral Data

In order to evaluate the proposed classification method in real analysis scenarios, we use the widely used hyperspectral data set collected by AVIRIS "AVIRIS Indian Pines Data Set". The scene contains 145 × 145 pixels and 220 spectral bands. The ground-truth data contains 16 mutually exclusive classes and a total of 10366 labeled pixels. This image constitutes a challenging problem due to the significant presence of classes with similar spectral signatures and also because of the unbalanced number of available labeled pixels per class.

To evaluate the performance of the proposed method, EMAP were built using threshold values in the range of 2,5% - 10% with a step of 2,5% for the standard deviation attribute and thresholds of 200, 500 and 1000 for the area attribute. In the training step, we used new data generated according to the interpolation of *l=50* training samples for each class. If the total number of labeled samples in the reference map for a single class L_c<50, we take *l=10*.

Table 1, illustrates the OA, AA, kappa statistic coefficient (k), and individual class accuracy (in percent) results achieved by the proposed classification method when we generated from *l* 50 new samples in each class and we used weighted summation kernel with μ=0.25. By adopting oversampling, the proposed method significantly improved the classification results obtained by the considered classification without oversampling.

For instance, the uses of linear interpolation for the spectral and spatial feature (Lin_Lin) obtained an OA of 93.21%, 4.39% larger than that obtained by SVM without oversampling. As a result, the obtained samples after oversampling play an important role in the classification process, it improves the accuracy of the supervised classifier (SVMs with composite kernel). It is remarkable that the uses of linear and cubic spline interpolation methods to oversample the available labeled samples leads to have more accurate classification than that resulted after the application of the other combinations between Lagrange, cubic spline and linear interpolation techniques. This indicate that the samples generated by the combination of linear and cubic spline interpolation are properly created refer to their similarity to the oversampled data in each class.

For illustrative purposes, Figure 10 shows the ground truth and some of the classification results obtained by the different methods for the AVIRIS Indian Pines scene. For each method, we randomly selected one of the maps obtained after conducting ten Monte Carlo runs. As shown by Figure 10, the classification after oversampling by Lin_Lin combination produced the best classification map. An immediate issue resulting from experiments in Figure 9 is whether the number of training samples increase (classification without oversampling uses 50 labeled samples while after oversampling the classifier use (50+50) labeled samples) could result in an improve in the classification results. In order to analyze this issue, we will increase the number of generated samples from *l=50* available learning data with using Lin_Lin combination in the oversampling step. Figure 11 (a) and (b) shows the OA and the kappa results as a function of generated data set size. From figure 10, we can conclude that the classification performance indeed depends on the size of generated training samples. In fact, when $50 \leq g \leq 200$, the increasing in the size of generated data increase the performance of the classification but the increasing also of g (g>200) decrease the accuracy of the classifier. This indicates that the increasing of g can lead to have bad training samples that confuse the classifier.

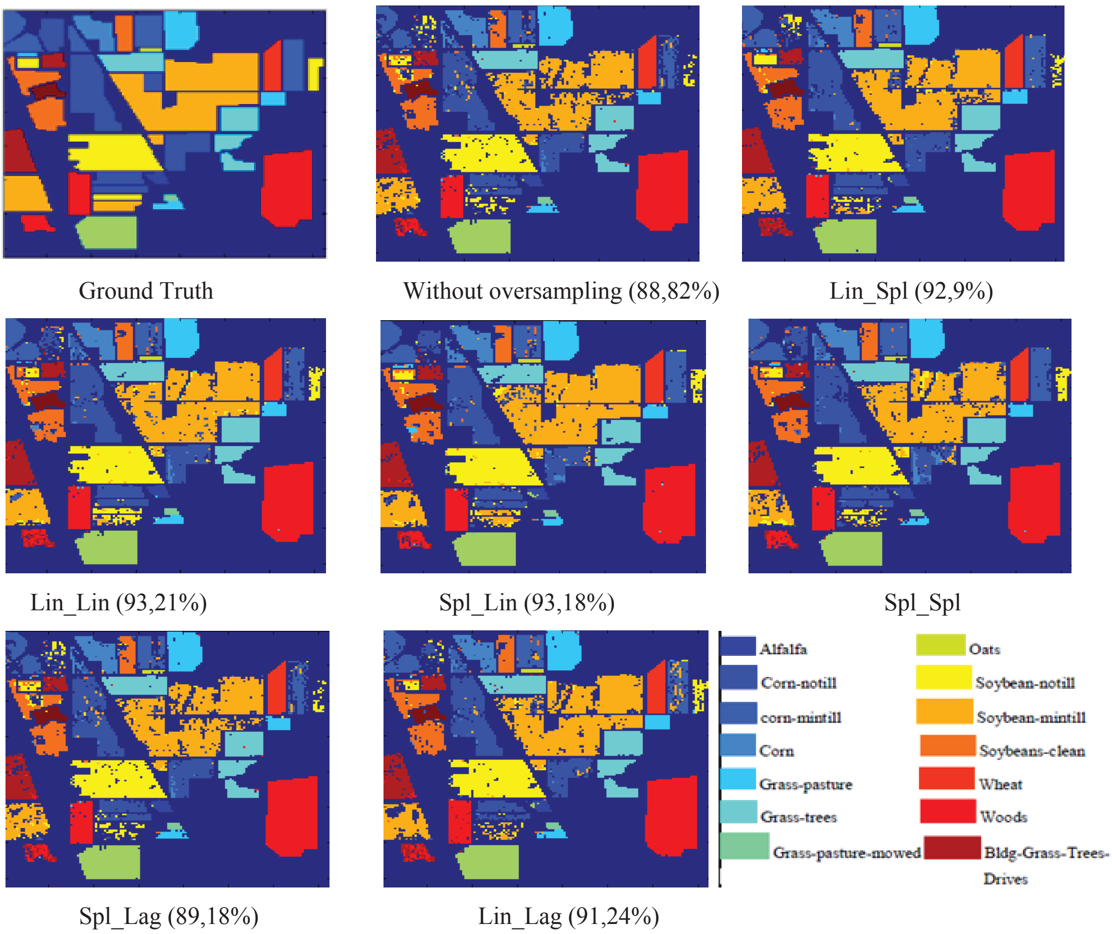
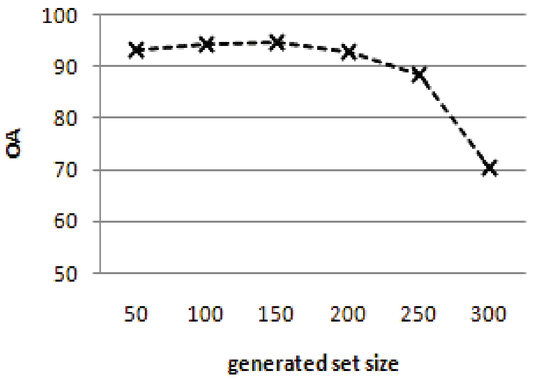


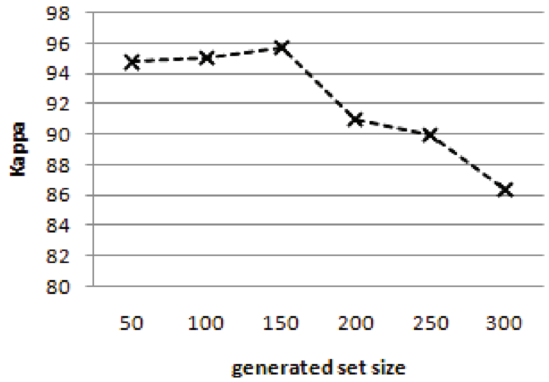
Figure 10. Classification maps obtained by the different interpolation methods combination for the AVIRIS Indian Pines scene (overall accuracies are reported in parentheses).

Table I: Overall, average, and individual class accuracies (in percent) and κ statistic obtained for the AVIRIS Indian Pines data set.

Class	Samples	Classification without oversampling	Classification after oversampling								
			Lin_Spl	Lin_Lin	Spl_Lin	Spl_Spl	Spl_Lag	Lag_Spl	Lag_Lin	Lin_Lag	Lag_Lag
Alfalfa	54	75	79,55	92,68	70,45	88,64	88,64	86,36	80,5	74,42	70,45
Corn-notill	1434	69,85	77,66	77,13	76,53	71,11	70,7	71,07	70,1	69,25	71,11
Corn-mintill	834	71,80	84,43	77,84	80,47	89,09	85,06	56,31	69,80	75,13	57,13
Corn	234	84,07	88,40	93,48	90	91,06	92	89,44	81,45	84,15	90,00
Grass/pasture	497	92,97	93,17	93,14	90,32	88,49	89,5	88,74	82,97	91,38	89,50
Grass/trees	747	95,03	97,96	96,38	92,65	98,26	93,26	95,92	94,65	95,80	92,75
Grass/pasture-mowed	26	87,50	93,75	93,33	87,5	87,50	87,50	93,33	87,50	93,75	87,50
Hay-windrowed	489	99,08	97,90	97,70	99,07	99,31	98,00	97,22	98,00	99,08	97,90
Oats	20	50	100	100	100	55,56	53,50	66,67	75,50	66,67	75,50
Soybean-no till	968	77,68	81,25	86,93	84,33	85,75	78,11	78,37	73,14	80,22	81,22
Soybean-min till	2468	69,40	79,77	81,27	84,56	74,23	74,23	67,88	67,88	76,18	71,80
Soybean-clean till	614	82,29	86,23	82,19	80,25	86,69	80,04	76,66	82,29	84,38	84,38
Wheat	212	100	100	99,38	99,37	99,38	99,37	98,76	100	100	99,38
Woods	1294	89,64	90,49	94,55	95,05	95,85	89,9	93,06	90,23	92,57	89,57
Bldg-Grass-Trees-Drives	380	88,07	93,87	94,17	88,07	84	87,00	80,73	88,07	87,20	88,20
Stone-Steel-Towers	95	100	100	100	100	95,56	81,11	97,73	96,30	86,36	86,36
OA		88,82	92,9	93,21	93,18	92,28	90,11	89,18	89,5	91,24	88,87
AA		83,27	90,27	91,26	88,66	86,9	84,24	83,64	83,76	84,79	83,3
kappa		93,5	94,55	94,75	93,04	92,49	91	93,7	93,5	94,45	93



(a)



(b)

Figure 11. Variation of OA (a) and kappa coefficient (b) according to the increase in generated set size.

5. Conclusion

In this paper, we have developed a new spectral-spatial supervised classification approach which combines spectral and spatial features via composite kernels and generate new training samples to overcome the problem of the limitation of labeled samples widely proposed in the classification of hyperspectral images. It investigates interpolation techniques to oversample the small set of training examples. By using weighted summation kernel with μ=0.25 for SVMs, the proposed method provides good accuracies when compared with the classification without oversampling. It also exhibits robustness to different criteria, such as the limited availability of training samples and the type of adopted kernel. Although our experimental results are competitive and encouraging when dealing with ill-posed problems, i.e., limited training samples versus high dimensionality of the input data, further work should be focused on oversampling step by applied another interpolation techniques.

6. References

[1] G. Camps-Valls, D.Tuia, L. Bruzzone and J. Atli Benediktsson," Advances in Hyperspectral Image Classification: Earth Monitoring with Statistical Learning Methods," IEEE Signal Processing Magazine, vol.31, no.1, pp. 45 - 54, Jan. 2014.

[2] S. Serpico and G. Moser, "Extraction of spectral channels from hyperspectral images for classification purposes," IEEE Trans. Geosci. Remote Sens., vol. 45, no. 2, pp. 484–495, Feb. 2007.

[3] M. Pal and G. M. Foody "Feature selection for classification of hyperspectral data by SVM", IEEE Transactions on Geoscience and Remote Sensing, vol. 48, pp. 2297-2307, 2010.

[4] M. Dalla Mura , A.Villa, J.A. Benediktsson and J. Chanussot "Classification of Hyperspectral Images by Using Extended Morphological Attribute Profiles and Independent Component Analysis", IEEE on geoscience and remote sensing letters, VOL. 8, NO. 3, pp. 542 - 546, May 2011.

[5] K. Tan, E. Li, Q. Du and P. Du "Hyperspectral Image Classification Using Band Selection and Morphological Profiles", IEEE journal of selected topics in applied earth observations and remote sensing, VOL. 7, NO. 1, pp. 40-48, JAN. 2014.

[6] X.Zhu, "Semi-supervised learning literature survey," comput. Sci. , Univ. Wisconsin-Madison, Madison, WI, Tech. Rep. 1530, 2005.

[7] M.Muniz-Mari, D.Tuia and G. Camps-Valls, " Semisupervised classification of remote sensing images with active queries," IEEE transactions on geoscience and remote sensing, VOL. 50, NO. 10, pp. 3751 - 3763, Oct. 2012.

[8] I. Dópido, J. Li, P. R. Marpu, A. Plaza, J. M. Bioucas Dias and J. A. Benediktsson, "Semisupervised Self-Learning for Hyperspectral Image Classification", IEEE transactions on geoscience and remote sensing, VOL. 51, NO. 7, pp. 4032 - 4044, JULY 2013.

[9] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. Vila-Frances, and J. Calpe-Maravilla "Composite Kernels for Hyperspectral Image Classification", IEEE geoscience and remote sensing letters, VOL.3, NO.1, 93 - 97, Jan. 2006.

[10] J. Li, P. R. Marpu, A. Plaza, J. M. Bioucas-Dias and J. A. Benediktsson " Generalized Composite Kernel Framework for Hyperspectral Image Classification", IEEE transactions on geoscience and remote sensing, VOL. 51, NO. 9, pp. 4816 - 4829, Sep. 2013.

[11] X. Huang and L. Zhang "An SVM Ensemble Approach Combining Spectral, Structural, and Semantic Features for the Classification of High-Resolution Remotely Sensed Imagery", IEEE transactions on geoscience and remote sensing, VOL. 51, NO. 1, pp. 257 - 272, Jan. 2013.

[12] R. Ben salem, K.S. Ettabaa and M. A. Hamdi "Spectral-spatial classification of hyperspectral images using different spatial features and composite kernels," Image Processing, Applications and Systems (IPAS), pp.1-7, Nov. 2014.

[13] P. Ghamisi, M. S. Couceiro, M. Fauvel and J. A. Benediktsson, "Integration of Segmentation Techniques for Classification of Hyperspectral Images", IEEE on geoscience and remote sensing letters, VOL. 11, NO. 1, pp. 342-346, JANUARY 2014.

[14] G. Mountrakis, J. Im, C. Ogole, "Support vector machines in remote sensing: A review," ISPRS

journal of photogrammetry and remote sensing, vol. 66, n. 3, pp. 247–259, May 2011.
[15] M. Ben salem, K.S. Ettabaa and M. A. Hamdi "Anomaly detection in hyperspectral imagery: an overview," Image Processing, Applications and Systems (IPAS), pp.1-6, Nov. 2014.
[16] K.S. Ettabaa, R. Ben salem and M. A. Hamdi " SVM for hyperspectral images classification based on 3D spectral signature," Advanced Technologies for Signal and Image Processing (ATSIP), pp. 42 - 47, March 2014.
[17] M.D. Mura, J.A.Benediktsson, B. Waske, L. Bruzzone, "Extended profiles with morphological attribute filters for the analyses of hyperspectral data," Int. J. Remote sens., vol. 31, N. 22, pp.5975-5991, Jul. 2010
[18] G. Mercier, M. Lennon, " Support vector machines for hyperspectral image classification with spectral based kernels," International symposium on geoscience and remote sensing (IGARSS 2003), vol.1, pp. 288 - 290 , July 2003.

Biographies

Rafika Ben Salem received the Eng. and M. degrees from National school of engineering, Sousse at 2009 and 2011. Currently she is a PhD student at National institute of applied sciences and technology, Tunis in the MMA Laboratory. Her research interests include remote sensing image interpretation and classification.

Karim Saheb Ettabaa received the M.E. and Dr. Eng. degrees from ENSI, Manouba, Tunisia, in 2004 and 2007, respectively. He is a Permanent Researcher at laboratory RIADI, National School of Computer Sciences Engineering and ITI, Telecom Bretagne, Brest Iroise Technopole CS 81828. His research interest includes image processing, data mining, artificial intelligence, pat- tern recognition, and their application to remote sensing

Mohamed Ali HAMDI received the Ph.D. degree in image processing, from the National University of engineering of TUNIS, He is Assistant Professor of Applied and Computational Mathematics and electronic, National institute of applied sciences and technology. His research interests are in the areas of digital signal processing (DSP), statistical estimation and their applications to signal and image processing and scientific computing.

www.icgst.com

Evaluation of the Taxonomic Consistency of Ontologies based on WordNet Hierarchical and Lexical Relations

Hafedh Nefzi, Mohamed Farah, Imed Riadh Farah
National School of Computer Science, University of Manouba, SIVIT-RIADI Laboratory, University Campus of Manouba 2010, Tunisia
hafedh.nefzi@gmail.com, mohamed.farah@riadi.rnu.tn, riadh.farah@ensi.rnu.tn,
http://www.riadi.rnu.tn/

Abstract

Presently, a considerable number of knowledge engineering researches have focused on the automatic building of ontologies. However, the uncertainty of the techniques and eventual heuristics adopted during the construction process has led researchers to explore methods for verifying and improving the quality of the outputs. In this intention, we propose a vision for checking the hierarchical structure of ontologies based on the WordNet lexical database as a background knowledge source. In order to test our work, we try to apply our proposed method on an existing valid geographic objects ontology.

Keywords: *geographic objects, knowledge modeling, ontology building, Taxonomic structure, similarity measures, evaluation, quality.*

1 Introduction

Satellite imagery is a relevant source of information for the identification of objects that make up the surface of the earth. Their exploitation in a spatio-temporal context helps to monitor and predict their behavior over time and to take appropriate decisions for the management of the environment.

Indeed, the advent of high-resolution images enabled the development of an object-oriented approach where the analysis of a scene is attached to groups of pixels representing concrete objects having a specific semantic. However, this progress leads to a large amount of available information that cannot be processed in its entirety by domain experts. This motivated the interest of research on the full or partial automation of the process of knowledge representation and extraction applied to satellite imagery. To use geographic image databases, the researchers used several knowledge representation formalisms, in particular the ontologies.

Nowadays, ontologies are becoming very popular in the area of knowledge management and sharing, especially after the evolution of the Semantic Web. They are considered as one of the most powerful tools for knowledge representation and reasoning. They aim to provide a commonly accepted understanding of a specific domain through the generic modeling, the exchange and the sharing of its specific knowledge. Knowledge is modeled in the form of concepts and their relations to each other. Several studies were interested in the use of standardized ontologies to share and annotate satellite image information [7, 1, 8, 4, 26, 20, 31]. The majority of these works presupposes the existence of a domain ontologies that may be developed, or be carried out, within the target application [4]. However, few studies have focused on their evaluation or validation.

In fact, the quality of an ontology is too sensitive to many parameters such as the consistency of the semantic resources from which it is built and the used techniques and heuristics to extract and organize relevant knowledge [19]. Therefore, as all engineering artifacts, assessing the quality of ontologies still remains an important issue for ontology engineering. The evaluation covers the structure and the content of ontologies and allows to verify several related criteria such as their consistency and their adequacy to the user's requirements and pre-established constraints.

In this paper, our main research question is how to examine taxonomic structure of a given geographic objects ontology. Firstly, we summarize the main evaluation alternatives. Secondly, we expose our method and the related structural measure for verifying the ontology hierarchical structure based on the WordNet[1] lexical database. Thirdly, we reserve the last section to an experimental study in which we expose and interpret the results of the application of our proposal on a geographic objects ontology.

[1] https://wordnet.princeton.edu/

2 Ontology evaluation: State of the art

Evaluation is a crucial phase in the building process of ontologies. It helps to simplify their development, to ensure their relevance to the requirement of a particular domain and to detect eventual ontology changes. However, the lack of unifying framework for methods and metrics for evaluating ontologies have led to several trials, each of which defines its own method and set of metrics. In this section, we try to summarize the main evaluation methods that can be classified according to their purpose into three categories: ranking, correctness, or quality.

When trying to reuse the already existing onotologies for a particular study domain, we are faced with the problem of determining the suitable ones for our needs. In this context [17] have presented an approach for clustering ontology. The main goal of this approach is to use a set of similarity measures for comparing ontology-based meta-data. Based on this work, [27] have developed the OntoQA approach that analyzes ontology schemas and their populations and describes them through a well defined set of schema and meta-data metrics. The first group includes the diagram metrics of ontologies, whose intention is to evaluate the ontology design and its potential for knowledge representation. The second group is interested in evaluating the structure of the knowledge base and more specifically how data is placed in ontology.

Further, the ranking category includes approaches for ranking and selecting ontologies. These approaches allow ranking a set of candidate ontologies in order to choose the most appropriate for a particular task. Ontometric [15] is one of the main used methods for systematic ontology selection, it aims to suggest the best ontology for a particular project on the basis of 160 properties organized in five dimensions of quantitative measurements: content, language, methodology, tool and costs. [21], have provided a corpus-based method to evaluate the functional adequacy of ontologies. [22] have proposed an ontology selection and ranking model consisting of selection standards and metrics based on better semantic matching capabilities. The proposed model allows to enhance the ontology selection and ranking method practically and effectively by enabling semantic matching of taxonomy or relational linkage between concepts and to identify what measures should be used to rank ontologies in a given context and what weight should be assigned to each selection measure. FOEval [3] is another model which presents two main features: first, it enables users to select from a set of proposed metrics, those which they help in the ontology evaluation process; and to assign weights to each one based on assumed impacts on this process. Second, it enables users to evaluate locally stored ontologies, and/or request search engines for available ontologies. The main goal of this model is to ease the ontology evaluation task, for users wishing to reuse available ontologies, enabling them to choose the most adequate ontology to their requirements. To evaluate and rank candidate ontologies, FOEval use a set of metrics that include: coverage, richness, detail-level, comprehensiveness, connectedness and computational efficiency.

The correctness category includes the approaches accounting for the formal correctness of the ontological knowledge and used primitives. In this category, the best known approach is Ontoclean [12] which is designed in order to justify the kinds of decisions that experienced ontology builders make and to explain the common mistakes of the inexperienced, as it analyses the intentional content of concepts. It is based on principles of rigidity, identity, unity and dependence. Based on this method, [5] have developed a framework which looks for taxonomic aspects such as circularity and redundancy, as well as errors in disjoint groups. [28] have developed another tool for evaluating real-world ontologies. [30] have proposed a tool that evaluates correctness, where an internal evaluation is performed, based on the correct usage of OWL primitives.

The third category addresses the evaluation of the global quality of ontology. Following this approach, the EvaLexon method [25] aims to evaluate the ontologies during their development from texts. It measures the most appropriate terms in ontology. The relevance of a term is judged by its frequency in the text from which the ontology was built and the list of terms for a specific domain. The evaluation is based on four metrics: precision, recall, coverage and accuracy. In turn, [9] have approached the ontology evaluation as a diagnostic task based on ontology descriptions, using three categories of criteria: structural (depth, breadth, tangledness, dispersion, consistency, anonymous classes, cycles, and density), functional (competence adequacy, functional modularity, precision, recall and accuracy), and usability profiling (documentation, efficiency, interfacing). By combining the different measurable criteria for each category, nine quality principles (qoods) are defined: cognitive ergonomics, transparency, integrity and computational efficiency, meta-level integrity, flexibility, expertise compliance, conformity with extension, integration and adaptation procedures, generic access and organizational ability.

To assess the quality of evolving ontologies, [16] have proposed a set of cohesion metrics that are considered as stable, where their results do not depend on

the semantic or structural ontology representation. In the same way, [6] proposed the Onto-Evoal approach which is based on an evaluation model to guide the management of inconsistencies by assessing the impact of proposed resolutions on the content and use of the ontology. This model defines a set of quantitative metrics allowing choosing a resolution that preserves the quality of the evolved ontology. Quality criteria considered in the proposed approach are: complexity, cohesion, taxonomy, abstraction, modularity, completeness and understanding. By referring to the work of [10] and [11], [29] presents a theoretical framework for assessing the quality of an ontology for the Web. The framework summarizes ontology evaluation methods in two dimensions: ontology quality criteria (accuracy, adaptability, clarity, completeness, computational efficiency, conciseness, consistency, and organizational fitness) and ontology aspects (vocabulary, syntax, structure, semantics, representation, and context). Building on the two large meta-properties of unity and simplicity, [2] have developed an evaluation methodology called OntoAbsolute that allows to assess the taxonomic and non-taxonomic relationships, analyzes the conceptual structure and evaluates the ontology as a whole.

3 Proposed evaluation method

Our method of analysis of the taxonomic consistency of ontologies is based on two key elements (1) the projection of the ontology to evaluate on WordNet, and (2) the checking of the conformity of its hierarchical links compared to those linking WordNet corresponding synsets.

WordNet is an on-line lexical database that lists, classifies and connects in various ways the semantic and lexical content of a number of languages such as English and French [18]. For each word of the language, WordNet offers a list of synsets (synonym set) that correspond to all its possible meanings. The synset is the building block upon which rests the entire system. It corresponds to a group of interchangeable words denoting one sense or a particular purpose. Different words and synsets are interconnected by a number of lexical relations as the hyponymy/hyperonymy, holonymy/meronymy and synonymy/antonymy. These relationships can be exploited to explore the exact meaning of a given word. Its third release [2] offers a number of 155287 words expressing 117659 different meanings (synset).

These values reveal the semantic richness of WordNet and enhances the utility of its use as a reference taxonomy in order to verify the structure of ontologies. However, its generic nature assign a special attention to the polysemy problems. Indeed, for a given concept identifier, WordNet has multiple possible nodes, each of which is part of a particular context and refers to a different signification. Consequently, the good location of a concept in WordNet returns to find the synset that reflects its exact meaning.

[2] http://wordnet.princeton.edu/wordnet/man/wnstats.7WN.html

3.1 Projection of ontology on WordNet

The aim of this step is to locate the concepts of our ontology in WordNet that serves as a reference support for the analysis and validation of the ontology taxonomic structure.

For doing this, we are led to find for each concept the corresponding WordNet synset. It is obvious that this treatment can not be limited to a simple term search of the concept identifier in WordNet. knowing that the same word can support multiple meanings. Therefore, to be able to map a given concept in WordNet, we need to distinguish, among all proposed synsets, the one that better corresponds. Our solution is to involve the context of the concept in its marking task in WordNet. The context of a concept is described by its identifier, labels, comments, neighborhood and properties.

The most appropriate synset for a given concept is the one that shares with it the maximum of knowledge in terms of neighborhood and textual descriptions.

Figure 1: Mapping between ontology concepts and WordNet synsets

Given the following:

$W(F,S)$ which defines the vocabulary admitted by WordNet corresponding to a set of pairs (F,S), where F is the form of a string on a finite alphabet and $S=\{s/F\}$ is the set of senses supported by F. s denotes an element of the set of meanings S (i.e. a synset).

Let the function $P(c,s_i)$ (Equation 1) defines the degree of knowledge sharing between the concept c and the synset s_i that denotes the synset number i of the identifier name of c. The relevant synset to a concept c must check this commitment:

$$s_k = relSyn(c) \Rightarrow P(c,s_k) > P(c,sj) \;\; \forall s_k \neq s_j \quad (1)$$

The function P is described by the following algorithm:

- $syn(c)$ a function that returns the synsets related to the identifier of the concept c.
- $lab(c)$ a function that returns the set of labels of the concept c.

- $com(c)$ a function that returns the comments associated with the concept c.
- $super(c)$ a function that returns the direct subsumer of the concept c.
- $w\_com(s)$ a function that returns the significant words included in the comments associated with the concept c.
- $w\_syn(s)$ a function that returns the set of synonyms words related to the synset s.
- $w\_gloss(s)$ a function that returns the significant words that compose the definition associated to the synset s.

  and

For extracting the relevant synset to the root concept of our ontology, we can proceed as follows :

- If the concept has a single synset in WordNet, it is then the corresponding synset.
- If the concept has labels, the sharing degree between it and a given synset is described by the intersection of their respective labels and synonyms.
- If the concept has comments, the sharing degree between it and a given synset is described by the intersection of their respective comments and definitions.
- Otherwise, the selection can be done manually (only for the root concept).

```
Input : c (a concept name identifier)
BEGIN
    if(| syn(c) |= 1)
      P = 1
    else if (| lab(c) |> 0)
      P =| lab(c) ∩ w_syn(s) | / | lab(c) |
    else if (| com(c) |> 0)
      P =| w_com(c) ∩ w_gloss(s) | / | w_com(c) |
    else return -1 (Cannot locate the root c in WordNet)
END.
```

However, the identification of the corresponding synset for a given non-root concept c is based on the computation of the distance that separates this synset to that associated with the closest subsumer of c in the ontology to be evaluated. We assume that the relevant synset s_k to a given concept is the one that is connected with the smallest number of subsumption links; among other synsets s_i of the same concept; to the corresponding synset of its subsumer. In this situation, the degree of knowledge sharing is described by the formula 2.

$$P(c, s_i) = \frac{1}{distance(s_i, relSyn(super(c))} \quad (2)$$

3.2 Validation of ontology taxonomic structure

Once the concepts of the ontology to be evaluated are mapped with the WordNet synsets, it is now possible to check the compatibility between the taxonomic structure of the ontology and that of corresponding synsets.

The hypothesis on which we base our assessment is that a given subsomption relationship between two concepts is considered valid only if their corresponding synsets are connected by the shortest hyperonymy path compared to those linking the synset of the subsuming concept to all synsets associated with the other concepts.

Several graph-theoretic measures can be used to calculate the proximity between two synsets in WordNet. They are mainly based on the number of edges that separate two nodes in a taxonomy. The most commonly used measures in literature are Rada [23], Leacock & Chodorow [14], Hirst & St-Onge [13] and Wu & Palmer [32]. Rada measure is considered as the most obvious way to evaluate the semantic similarity in a hierarchical ontology. It corresponds to the shortest path between two concepts in an ontology where only taxonomic links are considered, i.e. hyperonymy and hyponymy. Leacock & Chodorow measure is an extension of Rada which is in fact normalized by introducing a division by the maximum hierarchy depth of the involved concepts. Path measure adopts the same principle as the previous two measures by considering the inverse of the number of nodes along the shortest path between two nodes. As for Hirst & St-Onge measure, similarity between two concepts is determined by the minimum number of direction changes of the path between the two concepts. Indeed, depending on this measure, we distinguish four relation types between two concepts which are extra-strong, strong, medium and week. Wu & Palmer measure evaluates the similarity between two concepts as the distance of their most specific common subsumer to the root of the ontology divided by the shortest path between them.

4 Experimentation

In this section, we expose and interpret the results of applying our proposed evaluation method on a part of the ontology of AKTtiveSA[3]. This ontology deals with a number of geographical aspects of the knowledge infrastructure for humanitarian and disaster relief operations. It encompasses a wide variety of conceptualizations including terrain features, transport routes, rivers, shorelines, terrain elevation data, etc. [24]. The part to which we will limit our experimen-

[3] http://www.zaltys.net/ontology/AKTiveSAOntology.owl

tal study represents a hierarchy of 23 concepts modeling some Earth hydrographic objects (Table 2). Our scope of analysis will be restricted to concepts whose names appear in WordNet.

Figure 2: Taxonomic structure of the AKTiveSA ontology

As indicated above, the rapprochement between the concepts of the ontology to evaluate and the WordNet synsets may be supported by the texts associated with them, but also by their subsumers in both hierarchies. The concepts of our ontology lack any label. Table 1 shows the associated comments for each concept of the analyzed part.

Table 1: AKTiveSA concepts and relative comments

Concept	Comments
Body of water	Represents planetary structures that are part of the hydrosphere and that have a primary substance composition of a water.
Aquifer	An aquifer is an underground structure of water-bearing, permeable rock.
Reservoir	
Pond	A pond is a body of water smaller than a lake. However the difference between a pond and a lake is largely subjective. The term pond usually describes small bodies of water, generally smaller than one would require a boat to cross. Another definition is that a pond is a body of water where even its deepest areas are reached by sunlight.
Lake	A lake is a body of water surrounded by land.
Stream	A stream is a body of water with a detectable current, confined within a bed and banks. Stream is also an umbrella term used in the scientific community for all flowing natural waters.
River	A river is a large stream, which may also be a water way.
Canal	Canals are man-made waterways, usually connecting existing lakes, rivers, or oceans. Irrigation canals are man-made waterways for the delivery of water and preceded the use of transportation canals used by barges or narrowboats on smaller canals, and by ships on ship canals that connect to the ocean.
Creek	In British English and Indian English usage, a creek is a tidal water channel. Creeks may often dry to a muddy channel with little or no flow at low tide, but often with significant depth of water at high tide.
Spring	A spring is a point where groundwater flows out of the ground, and is thus where the aquifer surface meets the ground surface.
Ocean	A large body of water constituting a principal part of the hydrosphere.

The results of the evaluation of the structural proximity between each concept and each of its corresponding synsets are given in Table 2. For each of these concepts, we indicate the Path similarity between each of its related synsets and the synset that corresponds to its closest subsumer concept in the AKTiveSA ontology. The most relevant synset for a given concept is that having the highest Path similarity value (written in bold).

Table 2: The taxonomic proximity values between concepts and related synsets.

concept	#n#1	#n#2	#n#3	#n#4	#n#5	#n#6
Pond	**0.33**	-	-	-	-	-
Aquifer	**0.16**	-	-	-	-	-
Lake	**0.5**	0.11	0.11	-	-	-
Stream	**0.5**	0.09	0.08	0.11	0.08	-
Ocean	**0.5**	0.11	-	-	-	-
Reservoir	0.1	**0.5**	0.35	0.25	-	-
Canal	**0.25**	0.12	0.10	-	-	-
Creek	**0.5**	0.11	-	-	-	-
River	**0.5**	-	-	-	-	-
Spring	0.09	0.09	**0.14**	0.11	0.09	0.08

Excepting the case of the concept *Canal*, the confrontation of the results detailed in Table 2 with the definitions of the concepts (Table 1) and their relevant synsets (Table 4) proves the effectiveness of our approach to identify concepts in WordNet. We clearly notice that the definitions of the AKTiveSA concepts are highly compatible with the glosses related to found synsets. Furthermore, the localization of concepts in WordNet helps to better understand their contexts. The example in Table 3 reinforces this idea and shows that, among the three synsets related to the concept *Canal*, the first synset has the highest structural similarity compared to representative sysnets of the other concepts.

Table 3: The Path similarity between the synsets of *canal* and the other concept synsets

	b.water #n#1	aquifer #n#1	reservoir #n#2	pond #n#1	lake #n#1
canal#n#1	0.33	0.12	0.20	0.20	0.25
canal#n#2	0.14	0.10	0.11	0.11	0.12
canal#n#3	0.11	0.12	0.09	0.09	0.10

	stream #n#1	river #n#1	creek #n#1	spring #n#3	ocean #n#1
canal#n#1	0.25	0.20	0.20	0.12	0.25
canal#n#2	0.12	0.11	0.11	0.10	0.12
canal#n#3	0.10	0.09	0.09	0.12	0.10

On the other hand, by browsing through the glosses of the three synsets of *Canal*, it seems clearly that the definition "*long and narrow strip of water made for boats or for irrigation*" associated with the third synset is more appropriate to the context of geographic objects than the first synset. A simple computation of the terminological intersection between the concept comments and both associated synset glosses can reinforce this attitude and allows us to conclude that this definition is the closest to

the concept *Canal*. So, it is the third synset that will be considered rather than the first. However, this terminological likeness does not imply any semantic relatedness. Indeed, the calculation of the structural proximity between each of these two synsets and the other synsets related to ontology concepts (Table 3), shows us clearly that the synset number 1 is the closest of the ontology context and justifies the interest in favoring such measure to the terminological ones.

Applying our method of projection of the ontology on WordNet, the found correspondences (concept, synset) and related scores are described in Table 4. For each concept, we show the number of candidates WordNet synsets as well as the identifier and meaning of its relevant synset.

Table 4: The concepts of AKTiveSA ontology and their associated synset glosses

Concept	nb. syn.	Id Syn.	Gloss
Body of water	1	Body_of_water #n#1	the part of the earth's surface covered with water (such as a river or lake or ocean).
Aquifer	1	aquifer #n#1	underground bed or layer yielding ground water for wells and springs etc.
Reservoir	4	reservoir #n#2	lake used to store water for community use
Pond	1	pond #n#1	a small lake.
Lake	3	lake #n#1	a body of (usually fresh) water surrounded by land
Stream	5	stream #n#1	a natural body of running water flowing on or under the earth
River	1	river #n#1	a large natural stream of water (larger than a creek)
Canal	3	canal #n#1	an indistinct surface feature of Mars once thought to be a system of channels; they are now believed to be an optical illusion.
Creek	2	creek #n#1	a natural stream of water smaller than a river (and often a tributary of a river).
Spring	6	spring #n#3	a natural flow of ground water.
Ocean	2	ocean #n#1	a large body of water constituting a principal part of the hydrosphere.

Once the ontology concepts are located in WordNet, it is now possible to check the conformity of their hierarchical structure compared to that of the corresponding synsets. Table 5 shows; for each of the considered concepts; the corresponding synset and its nearest subsumer. A subsumption link for c_1 to c_2 (i.e. c_1 isA c_2) is considered valid only if the synset related to c_2 is the nearest subsumer to that of c_1 compared to synsets of all other concepts.

The six firsts lines of the Table 5 prove the accuracy of *isA* relationships for the pairs of concepts (Reservoir, Lake), (Lake, Body_Of_Water), (Stream, Body_Of_Water), (River, Stream), (Creek, Stream) and (Ocean, Body_Of_Water). Starting from the hypothesis that the used ontology is proven, these results confirm the interest of our taxonomic evaluation method.

Table 5: The found synsets and their most closer subsumers

Synset	Most closer subsumer	Path value
Reservoir#n#2	Lake#n#1	0.5
Lake#n#1	Body_Of_Water#n#1	0.5
Stream#n#1	Body_Of_Water#n#1	0.5
River#n#1	stream#n#1	0.5
Creek#n#1	stream#n#1	0.5
Ocean#n#1	Body_Of_Water#n#1	0.5
Pond#n#1	Lake#n#1	0.5
Canal#n#3	none	-
Aquifer#n#1	none	-
Spring#n#3	none	-

On another hand, the seventh line shows that the concept *Pond* can be considered a sub-concept of *Lake*. This can be very convincing if we admit that the *pond* is a small area of *still water*, especially one that is artificial[4].

The three last lines show that none of the concepts *Canal*, *Aquifer*, and *Spring* has a subsumer among all other concepts of the analyzed part. In fact, WordNet considers that *Canal* is a water course rather than body of water. On their part, the concepts *Aquifer*[5] and *Spring*[6] are seen as geological formation. This same meaning is shared by other dictionaries such as Oxford Dictionaries. These results are very useful in the sense that they can help us to revise and correct certain taxonomic relationships.

5 Conclusion

In this paper, we firstly presented a brief literature review of the main ontology evaluation methods with a special focus on the taxonomy aspects. We secondly present our method to check the hierarchical structure consistency of a given ontology. Our method is based on two main phases enabling respectively to (1) project the concepts of the ontology to evaluate on WordNet, and (2) to check the conformity of their hierarchical links compared to those linking WordNet corresponding synsets. To accomplish this task, we

[4] Oxford Dictionaries : http://www.oxforddictionaries.com/

[5] Oxford Dictionaries : Aquifer is "a layer of rock or soil that can absorb and hold water"

[6] Oxford Dictionaries : Spring is "a place where water comes naturally to the surface from under the ground"

have used the Path graph-theoretic measure that calculates the length of the shortest path between two given synsets in WordNet. The results of the experimentation that we have described in the last section show that the proposed method is promising.

As for further research directions, we envisage to incorporate this method in a quality driven building process that will allow us to use existing geographic objects ontological resources for the conceptual enrichment of a given core ontology dedicated to model satellite images content.

References

[1] Jesús Manuel Almendros-Jiménez, Jose A. Piedra, and Manuel Canton. An ontology-based modeling of an ocean satellite image retrieval system. In *IEEE International Geoscience & Remote Sensing Symposium, IGARSS 2010, July 25-30, 2010, Honolulu, Hawaii, USA, Proceedings*, pages 2237–2240, 2010.

[2] M. Amirhosseini and J. Salim. Ontoabsolute as a ontology evaluation methodology in analysis of the structural domains in upper, middle and lower level ontologies. In *International Conference on Semantic Technology and Information Retrieval (STAIR)*, 2011.

[3] Abderrazak Bachir Bouiadjra and Sidi Mohamed Benslimane. Foeval: Full ontology evaluation. pages 464–468. IEEE, 2011.

[4] Jean Charlet, Sylvie Szulman, Guy Pierra, Nadia Nadah, Henry Valery Teguiak, Nathalie Aussenac-Gilles, and Adeline Nazarenko. Dafoe: A multimodel and multimethod platform for building domain ontologies. In *Deuxièmes Journées Francophones sur les Ontologies*, page 66, 2008.

[5] Oscar Corcho, A. Gómez-Perez, R. González-Cabero, and Mari Carmen Suárez-Figueroa. Odeval: A tool for evaluating rdf(s), daml+oil and owl concept taxonomies. In *Conference on Artificial Intelligence Applications and Innovations*, pages 369–382, 2004.

[6] Rim Djedidi and Marie-Aude Aufaure. *ONTO-EVO^{a}L* an ontology evolution approach guided by pattern modeling and quality evaluation. In *Foundations of Information and Knowledge Systems, 6th International Symposium, FoIKS 2010, Sofia, Bulgaria, February 15-19, 2010. Proceedings*, pages 286–305, 2010.

[7] N. Durand, S. Derivaux, G. Forestier, C. Wemmert, P. Gancarski, O. Boussaid, and A. Puissant. Ontology–based object recognition for remote sensing image interpretation. In *IEEE International Conference on Tools with Artificial Intelligence*, pages 472–479, Greece, 2007.

[8] Imed Riadh Farah, Wassim Messaoudi, Karim saheb ettabÃća, and Basel Solaiman. Satellite image retrieval based on ontology merging. *Graphics, Vision and Image Processing GVIP*, 8(2):45–53, July 2008.

[9] Aldo Gangemi, Carola Catenacci, Massimiliano Ciaramita, and Jos Lehmann. Modelling ontology evaluation and validation. In *The Semantic Web: Research and Applications*, volume 4011 of *Lecture Notes in Computer Science*, pages 140–154. Springer, Heidelberg, DE, 2006.

[10] Asunción Gómez-Pérez. Evaluation of ontologies. *Int. J. Intell. Syst.*, 16(3):391–409, 2001.

[11] Thomas R. Gruber. A translation approach to portable ontology specifications. *Knowl. Acquis.*, 5(2):199–220, June 1993.

[12] N. Guarino. *IEEE Intelligent Systems*, chapter Towards a Formal Evaluation of Ontology Quality, pages 1541–1672. IEEE, 2004.

[13] G. Hirst and D. St-Onge. Lexical chains as representations of context for the detection and correction of malapropisms, 1997.

[14] C. Leacock and M. Chodorow. Combining local context and wordnet similarity for word sense identification. In *MIT Press*, pages 265–283, 1998.

[15] A. Lzano-Tello and A. Gomez Perez. Ontometric: A method to choose the appropriate ontology. *Journal of Database Management*, 2004.

[16] Yinglong Ma, Xinyv Ma, Shaohua Liu, and Beihong Jin. A proposal for stable semantic metrics based on evolving ontologies. In *First IITA International Joint Conference on Artificial Intelligence, Hainan Island, China, 25-26 April 2009*, pages 136–139, 2009.

[17] Alexander Maedche and Valentin Zacharias. Clustering ontology-based metadata in the semantic web. In *Proceedings of the 6th European Conference on Principles of Data Mining and Knowledge Discovery*, PKDD '02, pages 348–360, 2002.

[18] George A. Miller, Richard Beckwith, Christiane Fellbaum, Derek Gross, and Katherine Miller. Introduction to wordnet: an on-line lexical database. *International Journal of Lexicography*, 1993.

[19] Hafedh Nefzi, Mohamed Farah, Imed Riadh Farah, and Basel Solaiman. A critical analysis

of lifecycles and methods for ontology construction and evaluation. In *1st International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, pages 48–53, Sousse, Tunisia, March 2014.

[20] Hafedh Nefzi, Mohamed Farah, Imed Riadh Farah, and Basel Solaiman. A semi-automatic mapping selection in the ontology alignment process. In *Knowledge Engineering and Ontology Development (KEOD'14)*, Roma, Italia, 2014.

[21] Yael Netzer, David Gabay, Meni Adler, Yoav Goldberg, and Michael Elhadad. Ontology Evaluation through Text Classification. In *Advances in Web and Network Technologies, and Information Management*, pages 210–221. 2009.

[22] Jinsoo Park, Sunjoo Oh, and Joongho Ahn. Ontology selection ranking model for knowledge reuse. *Expert Syst. Appl.*, 38(5):5133–5144, May 2011.

[23] R. Rada, H. Mili, E. Bichnell, and M. Blettner. Development and application of a metric on semantic nets. In Man and Cybernetics, editors, *IEEE Transaction on Systems*, pages 17–30, 1989.

[24] Paul R. Smart, Alistair Russell, Nigel R. Shadbolt, m. c. schraefel, and Les Carr. Aktivesa: A technical demonstrator system for enhanced situation awareness. *Computer Journal*, 50(6):703–716, 2007.

[25] P. Spyns. EvaLexon: Assessing triples mined from texts. Technical Report 9, Star Lab, Brussels, Belgium, 2005.

[26] C.D. Suriyakala, V. Vedanarayanan, and A.Senthil Kumar. Intelligent satellite image processing using ontology. In *Circuits, Power and Computing Technologies (ICCPCT), 2013 International Conference on*, pages 1184–1187, 2013.

[27] Samir Tartir, I. Budak Arpinar, Michael Moore, Amit P. Sheth, and Boanerges Aleman-meza. Ontoqa: Metric-based ontology quality analysis. In *IEEE Workshop on Knowledge Acquisition from Distributed, Autonomous, Semantically Heterogeneous Data and Knowledge Sources*, 2005.

[28] Johanna Völker, Denny Vrandečić, York Sure, and Andreas Hotho. Aeon - an approach to the automatic evaluation of ontologies. *Appl. Ontol.*, 3(1-2):41–62, January 2008.

[29] D. Vrandecic. *Ontology Evaluation*. PhD thesis, University of Karlsruhe, 2010.

[30] Taowei David Wang and Bijan Parsia. Ontology performance profiling and model examination: First steps. In *Proceedings of the 6th International The Semantic Web and 2Nd Asian Conference on Asian Semantic Web Conference*, ISWC'07/ASWC'07, pages 595–608, 2007.

[31] Xiaolei Wang, Nengcheng Chen, Zeqiang Chen, Xunliang Yang, and Jizhen Li. Earth observation metadata ontology model for spatiotemporal-spectral semantic-enhanced satellite observation discovery: a case study of soil moisture monitoring. *GIScience & Remote Sensing*, 53(1):22–44, 2015.

[32] Z. Wu and M. Palmer. Verb semantics and lexical selection. In *proceeding of the 32nd Annual meeting of the Associations for Computational Linguistics*, pages 133–138, 1994.

Biographies

Hafedh Nefzi received his M.Sc. degree in computer sciences from the university of Jendouba, Tunisia, in 2012. Currently, he is working toward the Ph.D. degree with the Ecole Nationale des Sciences de l'Informatique (ENSI), University of Manouba, Tunisia. He is also a Permanent Researcher at Laboratory RIADI, University of Manouba, since 2011. His work is mainly related with machine learning and knowledge modeling applied to the remote sensing images.

Mohamed Farah received his M.S. and PhD degrees in Computer Science from the University of Paris-Dauphine IX, France, respectively in 2000 and 2006. Currently, he is working as Permanent Researcher at Laboratory RIADI, ENSI, Tunisia. His research interests include information science, information retrieval, image indexing and retrieval, operations research, artificial intelligence, databases, knowledge engineering and their applications to remote sensing.

Imed Riadh Farah received the M.sc. degree from the ISG Institute of Computer Sciences, Tunis, Tunisia, in 1995, and the Dr.Eng. degree from the ENSI, Tunisia, in 2003. After working as a Research Assistant (from 1996) and a Permanent Researcher at Laboratory RIADI, ENSI (since 1995), he has been an Associate Professor at the University of Manouba, since 2010. He has been an Associate Researcher in the Department ITI-Telecom Bretagne, Brest, France, since January 2009. His research interests include image processing, pattern recognition, artificial intelligence, data mining, and their application to remote sensing. Currently, he is the Director of the Higher Institute of Arts and Multimedia. Dr. Farah is a member of Arts-Pi Tunisia.

www.icgst.com

An Enhanced Character Segment Method in Image-based Email Classification

Mallikka Rajalingam, Valliappan Raman, Putra Sumari
School of Computer Sciences, Faculty of Engineering Computing & Science,
Universiti Sains Malaysia, Penang, 11800, Swinburne University of Technology, Sarawak, Malaysia.
mallikka2002@gmail.com, vallimaster@yahoo.com, putras@cs.usm.my

Abstract
As the number of internet users increases, email has been an efficient and popular communication mechanism. The integral of classification of text and image based email by machine learning approaches were used. Image email is one of the most recent tricks introduced by spammers. The unwanted text embedding into images which are sent as email attachments. This paper proposed an enhanced way to segment characters in image based email classification domain. A hybrid approach of Discrete Wavelet Transform (DWT) and Hough Transform based character segmentation is used. This algorithm can enhance the segmentation accuracy. The experimental results show that proposed method produces less number of false negatives when compared with existing machine learning techniques.

Keywords: *character segmentation, discrete wavelet transform, email classification, Hough transform.*

Nomenclature:

DWT	Discrete Wavelet Transform
OCR	Optical Character Recognition
CC	Connected Component
FP	False Positive
FN	False Negative
TP	True Positive
TN	True Negative

1. INTRODUCTION

Character segmentation is an active area in machine learning approach. In large and complex images, it is vital to segment the image and then recognize the characters using character segmentation method. Email message bodies consists of subject, header information and content of the email. Often spam images are constructed by introducing random changes to a given template image, to make signature-based detection techniques ineffective, and are obfuscated to prevent optical character recognition (OCR) tools from reading the embedded text.

Connected Component (CC) based methods begin with extraction and localize the text regions by processing only connected component information. There are three types of complications to be tackle in connected components, 1) to extract text 2) To filter out non-text objects 3) To infer text blocks from connected component. After the connected component extraction, CC-based approaches filter out non-text object. Figure 1 shows sample example of real spam image email [1].

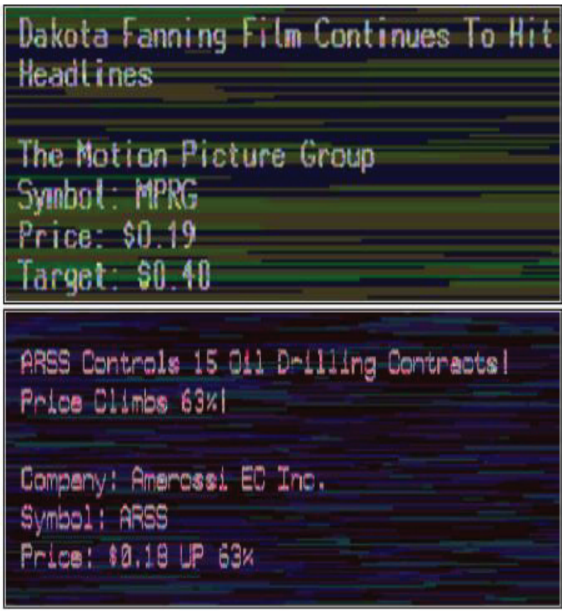


Figure 1. Example of spam image

Text line segmentation methods have been progressed within diverse projects which perform RGB to gray scale, gray scale to binary conversion, implemented hybrid approach and finally, input images are segmented as lines and characters. Otsu's thresholding method involves iterating through all the possible threshold values and calculating a measure of spread

for the pixel levels each side of the threshold, i.e. the pixels that fall in either foreground or background.

The rest of the research paper is organized as follows. Section 2 presents a survey of existing character segmentation algorithms in related work. We discuss character segmentation algorithm and methodology in section 3. We present and analyze the results of the experiment in section 4 and the conclusion with future work is described in section 5.

2. RELATED WORK

Researchers [2] presented a robust segmentation method for text extraction from the historical document images. The method is based on Markovian-Bayesian clustering on local graphs on both pixel and regional scales. It consists of three steps. In the first step, an over-segmented map of the input image is created. The resulting map provides a rich and accurate semi-mosaic fragments. The map is processed in the second step, similar and adjoining sub-regions are merged together to form accurate text shapes. The output of the second step, which contains accurate shapes, is processed in the final step in which, using clustering with fixed number of classes, the segmentation will be obtained. The method employs significantly the local and spatial correlation and coherence on both the image and between the stroke parts, and therefore is very robust with respect to the degradation. The resulting segmented text is smooth, and weak connections and loops are preserved thanks to robust nature of the method.

In the proposed work [3] presented a methodology for extracting text from images such as document images, scene images etc. Discrete wavelet transform (DWT) is used for extracting text information from complex images. For extracting text edges, the sobel edge detector is used. There are two different approaches have been used for text extraction from complex images namely region based approach and texture based approach. The region based method uses the properties of the color or gray scale in a text region or their differences regarding the background. This method is basically divided in two sub categories: edge based and connected component (CC) based methods. The edge based method is mainly focus on the high contrast between text and background. Connected component based method considers text as a set of separate connected components, each having distinct intensity and color distribution. The texture approach is based on the concept of textural properties. In this method, Fourier transforms. Discrete cosine transform and wavelet decomposition are generally used.

Researchers [4] proposed a hybrid approach which combines connected component analysis and an unsupervised thresholding for separation of text from the complex background. This proposed approach identifies the candidate text regions based on edge detection followed by a connected component analysis. Because of background complexity it is also possible that a non-text region may be identified as a text region. To overcome this problem extract texture features of connected components and analyze the feature values. Finally the threshold value for each detected text region is derived automatically from the data of corresponding image region to perform foreground separation. This proposed approach can handle document images with varying background of multiple colors. Also it can handle foreground text of any color, font and size. The experimental result shows that in high resolution document images, the false alarm rate is reduced and attain 97.8% of detecting the text region. In proposed work, fall through to detect single character without hole is the confinement need to improve.

A novel [5] hybrid approach of text segmentation using edge and texture feature information. The texture features such as homogeneity, contrast, energy for texts are different from non-text. The texture features are used to detect the text region from image. The edge based textures have many desired properties. The gradient magnitudes usually have higher values in the edge of the characters, even when the text is embedded in pictures.

Classification of [6] image spam and legitimate email using feed forward back propagation neural network (BPNN) model was proposed. In this technique the features of an image is extracted using histogram. In this method gradient histogram is extracted from an image. The obtained feature point of an image is then processed using Artificial Neural Network (ANN). In particular, this paper utilize feed forward back propagation neural network for classifying the image spam from those of legitimate mail ("ham").

3. PROPOSED CHARACTER SEGMENTATION ALGORITHM

SPAM EMAIL DETECTION METHOD:

1. **Get the combined words from the UI:**
   This can be done by converting string typed in mail, subject, and content into characters. So that we can get words to be compared with vocabulary dictionary
2. **Retrieve list of words from Vocabulary dictionary:**
   Try to create list of vocabulary with the words of spam filtering method
3. **Pre-processing email subject, mail ID, content**
   Using regular expression methods, we need to extract mail ids, subject and content with respect to SPAM vocabulary dictionary
4. **Tokenizing email**

Stemming the word by removing characters and separate out words and comparing with dictionary

5. **Matching with vocabulary list**
6. **If it matches with spam vocabulary classify as SPAM**
7. **Else it is HAM**

Segmentation is a process of partitioning a digital image into multiple segments. The objective of segmentation is to simplify and/or change the representation of an image into something that is more expressive and easier to analyze. This section explains how the image email characters are segmented. Figure 2 shows the detailed flow chart of character segmentation phase. It contains five stages. The stage 2 and 3 will be combined and also stage 4 and 5 are combined into single stage.

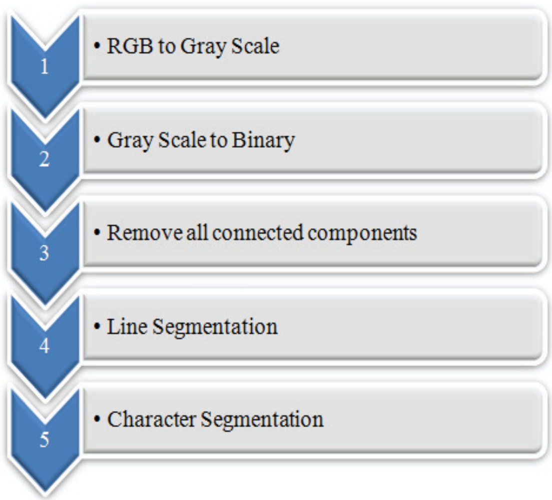


Figure 1. Detailed Flowchart of Character Segmentation

A. RGB to Gray Scale

In this process the color image is converted into gray scale image. To convert any color to a grayscale demonstration of its luminance, initially obtain the values of its red, green, and blue (RGB) primaries in linear intensity encoding, by gamma expansion. Then, add together 30% of the red value, 59% of the green value, and 11% of the blue value (these weights depend on the exact choice of the RGB primaries, but are typical). Regardless of the scale employed (0.0 to 1.0, 0 to 255, 0% to 100%, etc.), the resultant number is the desired linear luminance value; it typically needs to be gamma compressed to get back to a conventional gray scale representation. By human eye method,

Gray = 0.299 * Red + 0.587 * Green + 0.114 * Blue

To convert a gray intensity value to RGB, simply set all the three primary color components red, green and blue to the gray value, correcting to a different gamma if necessary.

B. Binarization

Binarization is performed so as to convert the RGB and gray scale images to the black and white pixel images. Only in a black and white image noise removal can be done efficiently without affecting the character pixels. The gray scale is converted into binary image using Otsu's method. Otsu's thresholding method involves iterating through all the possible threshold values and calculating a measure of spread for the pixel levels each side of the threshold, i.e. the pixels that fall in either foreground or background. The aim is to find the threshold value where the sum of foreground and background spreads is at its minimum.

After conversion of binarization removes all connected components (objects) that have fewer than 15 pixels from the binary image.

C. Line and Character Segmentation

After binarization the lines and characters are segmented. The hybrid approach Discrete Wavelet Transform (DWT) and Hough Transform is used to segment the characters.

DWT decompose signal into different components in the frequency domain, and the 2-d DWT in which it decomposes input image into four components or sub bands. One average component (LL) and three detail components (LH, HL, HH) as shown in figure 3, Sub bands are used to detect candidate text edges in the original image. The LL band is more significant band it contains more information of the original image, so it is most important part of the algorithm process. The LL sub-band can be further decomposed into four sub-bands. This process can continue to the required number of levels. It is known multi-level decomposition.

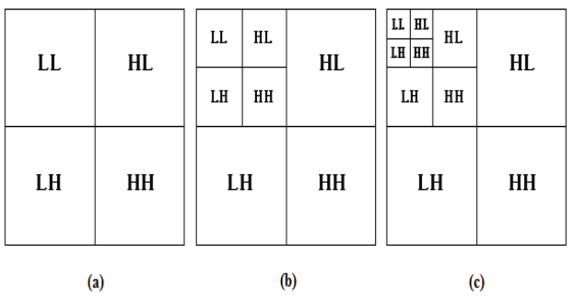


Figure 2. Different level of DWT Decomposition

The three level decomposition of the given digital image is as shown in figure 3. High pass and low pass filters are used to decompose the image first row-wise and then column wise.

The morphological operations like erosion and dilations are used for better approach of refining text region segmentation. Morphological operations are helpful in the removal of no texted regions. Various

types of boundaries like horizontal, vertical, diagonal etc. are clubbed together when they are segregated separately in unwanted non-text regions. The identified region of text consists of all these boundary and region information can be the area where such types of boundaries will be amalgamated. The boundaries are associated with one other in diversified directions and are normally short.

In most of the cases an edge detector can be used as a pre-processing stage to obtain image points or image pixels that are on the desired curve in the image space. But due to imperfections in either the image data or the edge detector there may be missing, isolated, or disjoint points or pixels on the desired curves as well as there may be spatial deviations between the ideal line or circle or ellipse and the noisy edge points as obtained from the edge detector. For these reasons, it is often non-trivial to group the extracted edge features to an appropriate set of lines, circles or ellipses. The purpose of the Hough transform is to address this type of problem by making it possible to perform groupings of edge points into object candidates by performing an explicit voting procedure over a set of parameterized image objects. Hough transform is applied on the binarized edge map DWT image to generate the Hough image of it. For this purpose, the parameters of the Hough transform, like theta (line rotation angle), rho (distance from the coordinate origin), peaks (row and column coordinates of Hough transform bins) are initialized or tuned in such a way that the lines are extracted as a set of connected words.

4. The Hough Transform Algorithm [7]

1. Find all of the desired feature points in the image space
2. For each feature point in image space
3. For each possibility *i* in the accumulator that passes through the feature point
4. Increment that position in the accumulator
5. Find local maxima in the accumulator
6. On requirement map each maxima in the accumulator back to image space

First the image has to be segmented row-wise (line segmentation), then each rows have to be segmented column-wise (character segmentation). The binary image is segmented row-wise (line segmentation).

5. RESULTS AND DISCUSSION

The proposed method was implemented using MATLAB (version R2013a) and the experiments are performed on an Intel(R) Core (TM) i5 machine with a speed 2.60 GHz and 8.0 GB RAM using Windows 8.1 64-bit Operating System. The resulted image of the line segmentation is shown in figure 4 and the resulted image of the Hough transform is shown in figure 5. The accuracy of classification is measured by computing the False Positive (FP), False Negative (FN), True Positive (TP), True Negative (TN), Recall, and Precision.

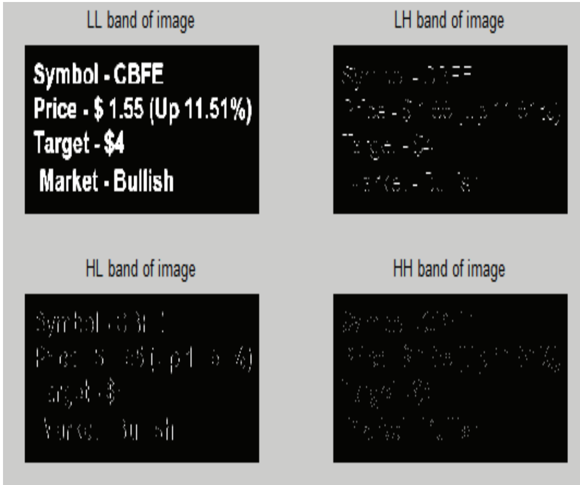


Figure 4. DWT of Image

Figure 5. Hough Transform

The outputs of each step in Character segmentation are illustrated in following figure 6, figure 7 and figure 8.

Figure 6. (a) Original Image, (b) Gray Scale Image, (c) Binary Image

Figure 3. Segmented Lines

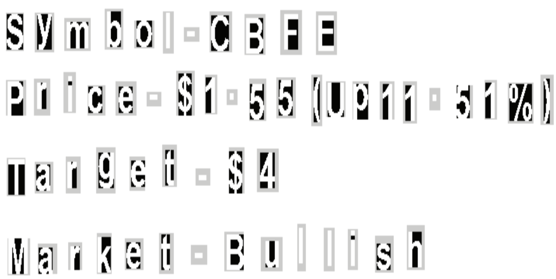


Figure 4. Segmented Characters

A combined-approach used for character segmentation i.e. Discrete Wavelet Transform (DWT) and Hough Transform techniques to achieve competitive segmentation accuracy. Morphological operations like erosion and dilations are used for better approach of refining text region segmentation. Morphological operations are helpful in the removal of no texted regions. The Hough transform technique is used for feature extraction in image analysis and most importantly designed for identification of lines in the image & identifying positions of arbitrary shapes. The purpose of Hough transform is to find imperfect or occurrence in the object within the shape. The output of DWT gives as input to Hough transform to get refined results. Table 1 shows the performance of image based email and result of total characters, extracted characters, correctly matched with given image email and incorrect match with given image emails.

TABLE I. PERFORMANCE OF IMAGE EMAIL

Image Email	Extracted Text	Result
	Symbol 1 CBFE Price 1 S 1l55 fUp 11l51Xj Tarqet 1 S4 Market l Bullish	Total Char - 70 Extracted Char - 70 Correct - 58 Incorrect - 12

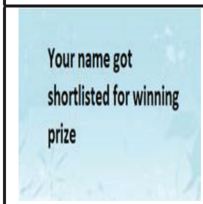
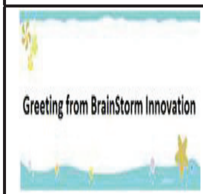
	Best prices on Sobare2 Windows XP f Ogioe XP z S80 f tons of other Best PRICED Sobare2 Cliok Here	Total Char - 103 Extracted Char- 98 Correct - 88 Incorrect - 10
	YAHOOz GROUPS	Total Char - 14 Extracted Char- 14 Correct Char - 13 Incorrect – 1
	Your name got shortlisted for winning 1 p rl z e	Total Char – 44 Extracted Char - 49 Correct - 39 Incorrect - 5
	Greeting from Brainstorm Innovation	Total Char - 36 Extracted Char- 36 Correct - 36 Incorrect - 0

Table II shows the image email accuracy from above table with image id for identification.

TABLE II. IMAGE EMAIL ACCURACY

Image Id	Accuracy
1	82.85 %
2	85.43%
3	92.85%
4	88.63%
5	100%

Table III shows the comparative study of existing methods merits and demerits. Proposed work shows the accuracy of 100% and eliminate the demerits of existing methods. Image based spam is a breakthrough from the spammers' view point; it is a simple and effective way of deceiving spam filters since they can process only text. The embedded image carries the target message and most email clients display the message in their entirety. Since many ham emails also have similar properties using HTML, carrying embedded images, with normal text as image-based

emails, existing spam filters can no longer distinguish between image-based spam and image ham.

TABLE III. MERITS AND DEMERITS OF EXISTING METHOD

Author & Year	Existing Method	Merits	Demerits	Accuracy in %
Jeffrey, Z. et al [8]	Character segmentation by Wavelet Transform	-Ability to detect edges in images -Less noise -High character segmentation	-In low image resolution, complexity is high	94.5
Sayal, N. et al [10]	Extract text from images by DWT, Gradient method and SVM classifier	-Remove non-text region	-Eliminating false negative (FN) rate is less	94.2
Singh, P. et al [9]	Character Extraction of colour image by DWT	-High Compression Ratio	-In reverse DWT, removing unwanted non text region is reduced	93.6
Saha, S. et al [7]	Text Segmentation by Hough transform	-Line extraction efficiently	-Extracting standard shapes -non uniform separation of some handwritten words	88%
Proposed work	Hybrid approach of DWT & Hough Transform	-Refine the character segmentation accuracy	✓	100%

For experimentation, Ling-Spam Corpus database are used for training and testing in the ratio of 70/30. The number of emails taken from ham and spam dataset using confusion matrix is shown in table IV.

TABLE IV. CONFUSION MATRIX OF EMAIL CLASSIFICATION

		Predicted	
		Ham	Spam
Actual	Ham	2310	102
	Spam	19	463

False Positives (FP) / False alarms are those regions in the image which are actually not characters of a text, but have been detected by the algorithm as text.

False Negatives (FN)/ Misses are those regions in the image which are actually text characters, but have not been detected by the algorithm

Recall rate (r) is defined as the ratio of the correctly detected characters to sum of correctly detected characters plus false negatives.

r = correctly detected characters / [correctly detected characters + FN]

Precision rate (p) is defined as the ratio of correctly detected characters to the sum of correctly detected characters plus false positives.
p = correctly detected characters / [correctly detected characters + FP]

F-score is the harmonic mean of the recall and precision rates.

The false negative rate is reduced based on the performance analysis, the results of TP, FP, TN, FN, precision, recall, F-measure, and overall accuracy are shown below:

TABLE V. TP, FP, TN, FN

Performance rate	
TPR	0.991842
FPR	0.180851
TNR	0.819149
FNR	0.008158

TABLE VI. RESULTS COMPARED WITH EXISTING METHOD

Author	Method/Algorithm	% of Accuracy
Dredze, M et al [11]	Feature based classification	89%
Battista Biggio et al [1]	VM	80%
R. Chandrasekaran et al [12]	SVM	93%
Meghali et al [13]	OCR, SVM	86%
Niti Syal et al [10]	DWT, Gradient Method and SVM	90%
Proposed work	DWT and Hough Transform	95%

Table VI shows result compared with some existing work. A hybrid approach of Discrete Wavelet Transform (DWT) and Hough Transform based character segmentation is applied in classification of spam email. This algorithm provided enhance the segmentation accuracy of 95%. The experimental results show that proposed method produces less number of false negatives when compared with existing machine learning techniques.

6. CONCLUSION WITH FUTURE WORK

The main objective of this paper explains about proposed algorithms, methods, and techniques used to segment characters in image based email. For experimentation, Ling-Spam Corpus database are used for training and testing. Image-based email performance evaluation with results are discussed and outputs are shown. The overall accuracy 100% is obtained by proposed hybrid approaches of character segmentation. The performance of false negative rate is reduced. To increase the size of training and testing data set will be the further enhancement of current work.

7. REFERENCE

[1] Biggio, b., fumera, g., pillai, i. & roli, f. (2011). "a survey and experimental evaluation of image spam filtering techniques. "elsevier. Pattern recognition. Pp. 1436-1446.

[2] HEDJAM, R., MOGHADDAM, R.F. & CHERIET, M. (2010). "TEXT EXTRACTION FROM DEGRADED DOCUMENT IMAGES." 2ND EUROPEAN WORKSHOP ON VISUAL INFORMATION PROCESSING (EUVIP). 2010. VOL.5. NO. 6. PP.247- 252.

[3] GUPTA, N. & BANGA, V.K. (2012). "IMAGE SEGMENTATION FOR TEXT EXTRACTION. " 2ND INTERNATIONAL CONFERENCE ON ELECTRICAL, ELECTRONICS AND CIVIL ENGINEERING (ICEECE, 2012). SINGAPORE. APRIL 28-29.

[4] SHIVANANDA, N. & NAGABHUSHAN, P. (2009). "SEPARATION OF FOREGROUND TEXT FROM COMPLEX BACKGROUND IN COLOR DOCUMENT IMAGES." SEVENTH INTERNATIONAL CONFERENCE ON ADVANCES IN PATTERN RECOGNITION. 2009. ICAPR '09. VOL.4. NO.6. PP.306-309.

[5] PATEL, P. & TIWARI, S. (2013). "TEXT SEGMENTATION FROM IMAGES."INTERNATIONAL JOURNAL OF COMPUTER APPLICATIONS. VOL. 67. NO.19.

[6] SORANAMAGESWARI, M. & MEENA, C. (2011). "A NOVEL APPROACH TOWARDS IMAGE SPAM CLASSIFICATION." INTERNATIONAL JOURNAL OF COMPUTER THEORY AND ENGINEERING. VOL.3. NO.1. PP. 84-88.

[7] SAHA, S., BASU, S., NASIPURI, M. & BASU, D. (2010). "A HOUGH TRANSFORM BASED TECHNIQUE FOR TEXT SEGMENTATION." JOURNAL OF COMPUTING. VOL. 2. NO. 2.

[8] JEFFREY, Z., RAMALINGAM, S. & BEKOOY, N. (2012). "REAL-TIME DSP-BASED LICENSE PLATE CHARACTER SEGMENTATION ALGORITHM USING 2D HAAR WAVELET TRANSFORM." ADVANCES IN WAVELET THEORY AND THEIR APPLICATIONS IN ENGINEERING, PHYSICS AND TECHNOLOGY. PP. 1-22.

[9] SINGH, P.N. & JAIN, A. (2014). "TEXT AND CHARACTER EXTRACTION OF COLOUR IMAGE USING DWT IN MATLAB IMAGE PROCESSING TOOL." "INTERNATIONAL JOURNAL OF ADVANCED TECHNOLOGY IN ENGINEERING AND SCIENCE. VOL.2, NO. 6. PP. 152-158.

[10] SYAL, N. & GARG, NK. (2014). "TEXT EXTRACTION IN IMAGES USING DWT, GRADIENT METHOD AND SVM CLASSIFIER." INTERNATIONAL JOURNAL OF EMERGING TECHNOLOGY AND ADVANCED ENGINEERING. VOL. 4. NO. 6. PP. 477-481.

[11] DREDZE, M., GEVARYAHU, R. & BACHRACH, A. (2007). "LEARNING FAST CLASSIFIERS FOR IMAGE SPAM." CEAS.

[12] R. CHANDRASEKARAN & RM. CHANDRASEKARAN (2011). "MORPHOLOGY BASED TEXT EXTRACTION IN IMAGES." INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY. VOL.2. NO.4. DECEMBER .PP. 103-107.

[13] DAS, M & PRASAD, V. (2014). "ANALYSIS OF AN IMAGE SPAM IN EMAIL BASED ON CONTENT ANALYSIS." INTERNATIONAL JOURNAL ON NATURAL LANGUAGE COMPUTING (IJNLC). VOL. 3. NO. 3. PP. 129-140.

Biographies

Mallikka Rajalingam received her M.Sc Information Technology from Bharathidasan University, Tiruchirappalli, India in 2005, M.Phil Computer Science from Madurai Kamaraj University, Madurai, India in 2008, M.Tech Computer Science & Engineering from SASTRA University, Thanjavur, India in 2009. She is currently pursuing the Ph.D. degree at the School of Computer Science, Universiti Sains Malaysia, Malaysia. She is the candidate of the Multimedia Computing Research Group, School of Computer Science, USM. Her research interests include image processing, computer vision, pattern recognition, character recognition, document image analysis, text analysis and multimedia networking.

Valliappan Raman is currently a senior lecturer in Faculty of Engineering Computing & Science at Swinburne University of Technology, Sarawak, Malaysia. He received B.E computer science from Madurai Kamaraj University, Madurai, India in 2005. M.Sc. computer science from Universiti Sains Malaysia, Malaysia. Ph.D. from Universiti Sains Malaysia, Malaysia in 2015. His specialization is in medical imaging and health informatics, computer vision and pattern recognition. He has published many papers in impact factor journals and well reputed conferences.

Putra Sumari is currently a associate professor in the School of Computer Science at Universiti Sains Malaysia, Penang, Malaysia. He received his B.Sc computer science (Hons) from Universiti Sains Malaysia, Malaysia in 1994. M.Sc video storage and retrieval from Liverpool University, England in 1997. Ph.D. from Liverpool John Moores Univeristy, England in 2000. His research interests include multimedia tools and applications, multimedia on demand & MPEG technology, data broadcasting, image/video retrieval, image/video compression and image/video watermarking.